NBSIR 88-3725

Evaluation of A Copy Prevention Method for Digital Audio Tape Systems

NATIONAL ENGINEERING LABORATORY
NATIONAL BUREAU OF STANDARDS

B.A. Bell and G.N. Stenbakken
Electronic Instrumentation and Metrology Group

D.R. Flynn, D.J. Evans, E.D. Burnett, and V. Nedzelnitsky
Acoustic Measurements Group

K.R. Eberhardt
Statistical Engineering Division

February 1988

U.S. Department of Commerce
C. William Verity , Secretary

National Bureau of Standards,
Ernest Ambler, Director

EXECUTIVE SUMMARY

Digital audio tape (DAT) systems have all the attractive audio qualities now associated with compact disc technology plus the ability to make very high quality copies of copies. Concern over loss of intellectual property rights has produced proposals for legislation requiring that DAT systems sold in the United States be fitted with a copy prevention decoder such that suitably coded material could not be copied by the DAT recorder. The CBS Records system, the basis for the proposals, removes a narrow band of frequencies from the audio signal - termed putting a "notch" in the frequency spectrum of the signal.

In considering this proposal, several committees of the Congress asked the National Bureau of Standards (NBS) three questions about the copy prevention system:

1. Does the copy prevention system achieve its purpose to prevent digital audio taping machines from recording?

2. Does the system diminish the quality of the prerecorded material into which the notch is inserted?

3. Can the system be bypassed, and, if so, how easily?

The National Bureau of Standards received from CBS Records descriptive material, specifications, circuit diagrams, and operating encoding and recording/decoding devices for the tests. The results and conclusions of the NBS tests apply to these devices. Suggestions for recorded material to be used in listening tests were received from several sources.

NBS conducted a series of laboratory studies to understand the copy prevention system and to exercise it in several ways to obtain answers to question 1 above. The NBS also constructed and evaluated several circuits designed to bypass the system (question 3). With the aid of an expert in acoustical evaluations on contract from a university, listening studies were conducted using subjects drawn from the local section of the Audio Engineering Society, NBS staff, and a few local audiophiles and musicians. These listening studies were carried out to answer question 2 above.

The NBS findings are as follows:

1. Answer to question 1 - "Does the system achieve its purpose?":

<u>NBS Conclusion</u>: The system does not achieve its stated purpose.

The system does prevent the copying of notched material much of the time. However, for about half of the recorded tracks studied, the system exhibited <u>false negatives</u>; i.e., notched material was nonetheless recorded. In addition, the system also exhibited <u>false positive</u> behavior, i.e., the system failed to record unnotched material. NBS studied 502 tracks on 54 compact discs and found false positives for 16 tracks on 10 discs.

2. Answer to question 2 - "Does the system diminish the quality of the prerecorded material?":

NBS Conclusion: The system's encoder alters the original electrical signal. For some listeners for some selections, this results in a discernible difference between prerecorded notched and unnotched material.

NBS interpreted this question both in terms of objective electrical measurements and in terms of whether or not listeners can detect a difference between notched and unnotched material in carefully controlled listening tests. Electrical measurements showed that the encoded signal is degraded relative to the signal as it was prior to copy prevention coding, degraded in terms of adding a deep notch in signal amplitudes near 3840 Hz, and degraded in terms of changes in signal phase relationships near 3840 Hz. In a double-blind series of listening tests, 87 subjects listened to a prerecorded tape of short selections from 24 different sources - 20 compact discs and four selections produced on a keyboard synthesizer. The results show that, although the effects of the encoder are fairly subtle for some musical selections, there are some selections for which the subjects detected differences between notched and unnotched material.

In a second series, 15 experienced listeners worked with 10 selections presented on parallel tape tracks such that the subject could switch back and forth from notched to unnotched material. For 2 of the 10 selections the encoded version was correctly identified 12 out of 15 times; these results are statistically significant.

3. Answer to question 3 - "Can the system be bypassed, and, if so, how easily?":

NBS Conclusion: The copy prevention system can be bypassed easily.

This question was interpreted by NBS as asking whether or not a DAT recorder can be made to record notched prerecorded material. It is not possible to restore all of the actual information removed by the encoding process. NBS engineers designed and constructed several electronic circuits for implementing five different methods to circumvent or defeat the copy prevention system by the use of external signal conditioning. All five methods succeeded in bypassing the copy prevention system. The circuits are simple and easy to construct. The cost would be of the order of $100.

Table of Contents

List of Figures

List of Tables

1. INTRODUCTION

Digital encoding of sound is causing a revolution in the world's audio industry. Digital sound systems have lower noise, less distortion, and a more even frequency response than conventional analog sound systems. Moreover, successive rerecordings of digital material produce negligible degradation of quality from generation to generation. The compact disc (CD) player is the first mass marketed hi-fi product to exploit this digital technology. The CD player commands a rapidly expanding market share. However, only playback versions are currently available. This situation is about to change with the introduction of digital audio tape (DAT) recorder/players. These promise to have all the qualities of the CD systems plus the ability to make very high quality copies of copies.

There is a concern that the quality of rerecording from DAT systems may cause financial harm to recording artists and companies by making it relatively easy for copies to be made and sold or used by others who might otherwise buy additional tapes and discs. This situation is seen as a violation of intellectual property rights. Not everyone agrees; there are those who feel that home copying for personal use should be permitted and only copying for profit should be barred. The controversy is similar to the one that occurred a few years ago over video taping and rerecording.

The Administration proposed in 1987 that Congress enact legislation requiring that DAT systems sold in the United States be fitted with a copy prevention decoder such that suitably "coded" material could not be copied by the DAT recorder. Bills were introduced in the House (H.R. 1384) and the Senate (S. 506). In the CBS Records system, which is the basis for the proposals, the "code" is achieved by removing a narrow band of frequencies from the audio signal - referred to as putting a "notch" in the signal's frequency spectrum. The copy prevention is accomplished by adding to the DAT system a decoder circuit for detecting the notch and shutting down the recorder.

Hearings were held at which the opposition to the proposal claimed that serious degradation of recorded sound would result from notching by the CBS method. Opponents also say that the CBS system may produce false positives, i.e., refusals to record unnotched material, and that the copy prevention system could in any case be easily defeated or circumvented by a determined pirate and even by home electronics buffs. The proponents strongly disagree.

The responsible Congressional committees decided to ask the National Bureau of Standards (NBS) to study the CBS copy prevention system and provide answers to the following questions (see Appendix A):

1. Does the copy prevention system achieve its purpose to prevent digital audio taping machines from recording?

2. Does the system diminish the quality of the prerecorded material into which the notch is inserted?

3. Can the system be bypassed, and if so, how easily?

The two groups representing the opposing views - the Home Recording Rights Coalition and the Recording Industry Association of America, Inc. - agreed to fund the NBS study. The Bureau requested descriptive material, specifications, circuit diagrams, and operating devices. CBS Records agreed to supply these items but officially requested that the Bureau not reveal information identified by CBS as company proprietary. NBS notified both parties that NBS would hold confidential information (the schematic diagrams and detailed specifications) supplied and identified as proprietary by CBS, but would publish all information developed by the NBS staff in the course of the study. Inevitably, the NBS measurements describe some of the subjects included in the CBS information. The Bureau also stated that it planned to subcontract the necessary listening studies to a university or firm specializing in psychoacoustic measurements.

The reader should keep firmly in mind that the National Bureau of Standards agreed to evaluate a specific implementation of the CBS Records copy prevention system. The Bureau did not undertake to study variations in (statistical) lots of devices supplied by CBS nor did NBS undertake efforts to improve upon the devices and information provided by CBS. In short, NBS worked with what was provided. Altogether, CBS sent two recorders/decoders and two encoders to NBS. The first recorder/decoder was defective at the outset and so the second unit was shipped. The second encoder was requested by NBS to speed up the evaluation. As will be discussed below, differences between the two encoders were noted but not addressed.

This report has been written to provide the answers to the three questions listed above and to supply the necessary supporting documentation of the laboratory work, while at the same time revealing a minimum of information identified by CBS as company proprietary. This situation may make complete understanding of some aspects a little more difficult than would otherwise be the case.

NOTE: Certain commercial equipment, instruments, or components are identified in this report in order to specify adequately the experimental procedure. Such identification does not imply recommendation or endorsement by the National Bureau of Standards, nor does it imply that the equipment or components are necessarily the best available for the purpose.

2. DESCRIPTION AND CHARACTERIZATION OF THE COPY-PREVENTION SYSTEM

CBS Records made available to the National Bureau of Standards (NBS) two copy-prevention encoders that were designed and built for use by recording companies to encode, or "notch," recorded material and two Sony rotary-head digital audio tape recorders (R-DAT), each equipped with a decoder that was designed and built by CBS Records to detect the presence of encoding.

Sections 2.1 and 2.2 describe the operation of the encoder and the physical measurements that were made in order to characterize its performance. In summary, the encoder consists of a narrow band-reject filter that removes signal components near a nominal center frequency of 3840 Hz. The encoder was found to be dynamic, with logic circuitry that determines whether to encode or not encode, depending upon the signal levels near 3840 and 2715 Hz.

Sections 2.3 and 2.4 describe the operation of the decoder and the physical measurements that were made in order to characterize its performance. In summary, the decoder scans the input signal components that pass through filters having center frequencies near 3300, 3800, and 4300 Hz, with these center frequencies being varied by about ±5 percent, to determine whether or not the signal is encoded. Since these frequencies vary in a manner not correlated with the incoming signal, the behavior of the decoder/recorder will not be exactly reproducible upon repeated tests with a given passage of recorded material, thus requiring a **statistical** approach to testing these circuits.

2.1 Encoder: Operation

The encoders had the serial numbers CCE-006 and CCE-012. In this report, the first encoder supplied by CBS Records is hereafter referred to as "Encoder 006" and the second encoder sent to NBS is hereafter referred to as "Encoder 012." When it is not necessary to distinquish between the two specimens of encoder that were provided, they are referred to simply as "the Encoder."

The Encoder provided to NBS is an analog device that filters the signal in each stereo channel so as to remove energy over a narrow band of frequencies. The front panel controls on the Encoder include a switch that enables the Encoder to be totally bypassed and another switch that disables the action of the Encoder but causes the signals for the two channels to pass through the unity gain amplifiers of the Encoder.

Conceptually, the Encoder operates as shown in figure 1. The signal for each channel passes through unity gain amplifiers (not shown) to a summing circuit ("SUMMER," at the top and bottom of figure 1). In parallel, each signal passes through a narrow bandpass filter and then to a solid-state switch that can be enabled to direct the filtered signal for each channel to the corresponding summing circuit. If the switch is "closed," each summing circuit subtracts the bandpass-filtered signal from its corresponding direct signal and passes the resultant signal on to the output of the Encoder. Thus, when the solid-state switches are closed, the Encoder acts as a dual-channel notch filter that removes energy from the audio signals over the range of

frequencies transmitted by the bandpass filters. This pair of notch filters is hereafter referred to in the singular as the encoding notch filter or simply as the encoding filter. When the solid-state switches are open, the signal passes through the Encoder without being "notched."

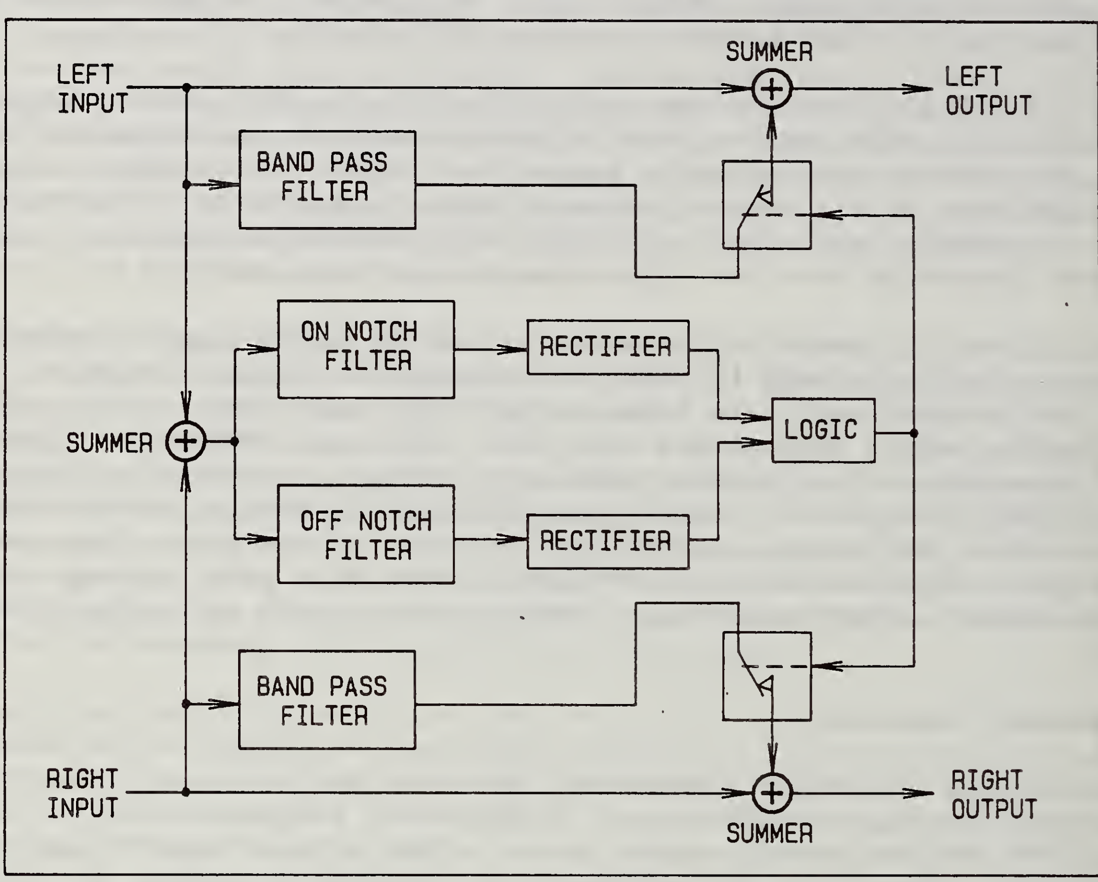


Figure 1 Conceptual block diagram of the major elements contained in the Encoder, as determined in the NBS study.

The operation of the solid-state switches shown in figure 1 is controlled by additional filters and logic circuits that operate as follows. The signals from the left and right channels are summed by a third summing circuit, shown at the center left of figure 1. This monaural signal is passed through a bandpass filter, hereafter referred to as the "on-notch filter," with a center frequency that is nominally the same as that of the encoding notch filter. The monaural signal is also passed through another band pass filter, hereafter referred to as the "off-notch filter," with a center frequency nominally one-half octave below that of the encoding notch filter. The outputs of the on-notch filter and the off-notch filter are each rectified, detected, and passed on to logic circuitry that activates or deactivates the two-channel encoding notch filter, depending upon the signal levels from the on-notch and off-notch filters. Note that the on-notch and off-notch filters never operate directly upon the signals going to the output of the Encoder.

The encoding notch filter removes energy at frequencies in the vicinity of 3840 Hz. In musical notation, this frequency falls between B-flat$_7$ and B$_7$, the highest B-flat and B on the standard piano keyboard, in the equally-tempered chromatic scale. (The subscript notation used is that which assigns to C$_0$ a frequency which corresponds roughly to the lowest audible pitch; in this scheme, the seventh octave extends from C$_7$ with a frequency of 2093.0 Hz through B$_7$ with a frequency of 3951.1 Hz. The next note above B$_7$ is C$_8$, which is four octaves above C$_4$, commonly known as middle C.) The level reduction near the center frequency of 3840 Hz is more than 80 dB, corresponding to a reduction to less than 0.01 percent of the original amplitude. At this degree of reduction, the signal effectively may be considered to be totally removed. At frequencies away from the center of the encoding notch filter, the reduction is less, becoming 3 dB (a 3-dB decrease represents a reduction in amplitude to 71 percent of the original amplitude) at about 110 Hz above and below the center frequency, i.e., at nominally 3730 and 3950 Hz. These two frequencies are very close to those of the musical notes, B-flat$_7$ and B$_7$. The off-notch filter has a center frequency that is near to 2715 Hz, the geometrical mean of the frequencies corresponding to the musical notes E$_7$ and F$_7$.

It should be noted that although 3840 Hz is a high frequency in terms of the musical scale, it is not at all above the region of significance to hearing. In fact, for young adults with normal hearing, the ear is most sensitive at frequencies near 3840 Hz.

Not many musical instruments are capable of producing notes that have fundamental frequencies falling at or above the center frequency of the encoding notch filter. However, when performing this study, we did observe fundamental notes in the encoding frequency region for piano, piccolo, and bells. Many organs and synthesizers can also produce such fundamental notes. Frequencies near 3840 Hz have significance for the perception of musical timbre because overtones of notes falling lower on the musical scale may occur in this region. Also, transient sounds, such as those produced by percussion instruments, will have energy over a broad range of frequencies, frequently including those in the encoding region. Instruments played slightly off pitch, or using a scale based on notes other than the 12 tones of the "western" scale will have many components in the encoding region. Synthesizers can produce notes of any frequency whatsoever, and several examples of components with frequencies "sliding" though the encoding region were noted. The use of vibrato, or the production of trills, with voice or instruments, especially stringed instruments, will cause energy which otherwise would be below or above the center frequency of the encoding notch filter (e.g., at the notes B-flat$_7$ and B$_7$) to shift into the encoding region. This effect was frequently observed in this study.

As indicated above, the Encoder contains logic circuitry which switches the encoding notch filter in and out of the circuit, depending upon the signal levels through the on-notch and off-notch filters. Presumably, this interruption of the notching operation is effected to prevent notching of audio signals when the encoding might be more noticeable. The results of NBS measurements of the influence of the on-notch and off-notch signal levels on the encoding operation are presented in section 2.2. However, in order to describe the operation of the Encoder, the results of these measurements are

qualitatively described here. The default operation of the Encoder is to engage the encoding notch filter, thus encoding both audio channels. However, if the off-notch signal is below a certain critical level and the on-notch signal is above a particular critical level, the logic circuitry will cause the Encoder to remove the encoding notch filter from both channels. When the off-notch signal and the on-notch signal are each above their respective critical levels, the action of the Encoder is dependent upon both of these signal levels.

In summary, the Encoder places a notch in both audio channels, with the center frequency of the notch occurring approximately midway between the musical notes of B-flat$_7$ and B$_7$, with there being some attenuation of the signal at these notes. Encoding does not occur at all times -- an attempt is made in the Encoder to eliminate encoding when it would produce effects that are more likely to be noticed.

2.2 Encoder: Physical Measurements.

This section describes the physical measurements that were made in order to characterize the Encoder and obtain quantitative information concerning its operation and the effects of the Encoder on the signals passing through it.

Encoding Notch Filter

The frequency characteristics, both amplitude and phase, of both channels of the encoding notch filter were measured for both Encoder 006 and Encoder 012 using a Hewlett-Packard (HP) Model 3562A Dynamic Signal Analyzer. Because of the dynamic nature of the Encoder -- the encoding notch filter switching in and out depending upon the signal levels from the on-notch and off-notch filters -- special steps had to be taken to force the encoding notch filter to be switched in while its physical characteristics were being measured.

Figures 2 and 3 show the amplitude-frequency characteristic and the phase characteristic, respectively, of the encoding notch filter for the left channel of Encoder 006. These data, shown over the frequency range from 0 to 5 kHz, were obtained using, as an input signal, noise which was pre-shaped to have a dip in its spectrum in the region around 3840 Hz, thus causing encoding to occur at all times. The amplitude-frequency characteristic is seen to be quite flat except for the deep notch centered nominally near 3840 Hz; cursors are included at the frequencies corresponding to the musical notes B-flat$_7$ and B$_7$, so as to help the reader to follow the different frequency scales used for the abscissa in subsequent plots. The phase characteristic is seen to deviate considerably from an ideal flat characteristic in the region around 3840 Hz. Note that this phase plot, as well as others to follow, "wraps" when the phase passes through ±180 deg, e.g., a phase of +200 deg is shown plotted at 200 - 360 = -160 deg.

The dynamic range of the analyzer is about 80 dB for this type of measurement using a random noise signal. Since the encoding filter introduces more than 80 dB attenuation near its center frequency, methods using random noise as the source may not give an accurate description of the filter characteristics in the region of its maximum attenuation.

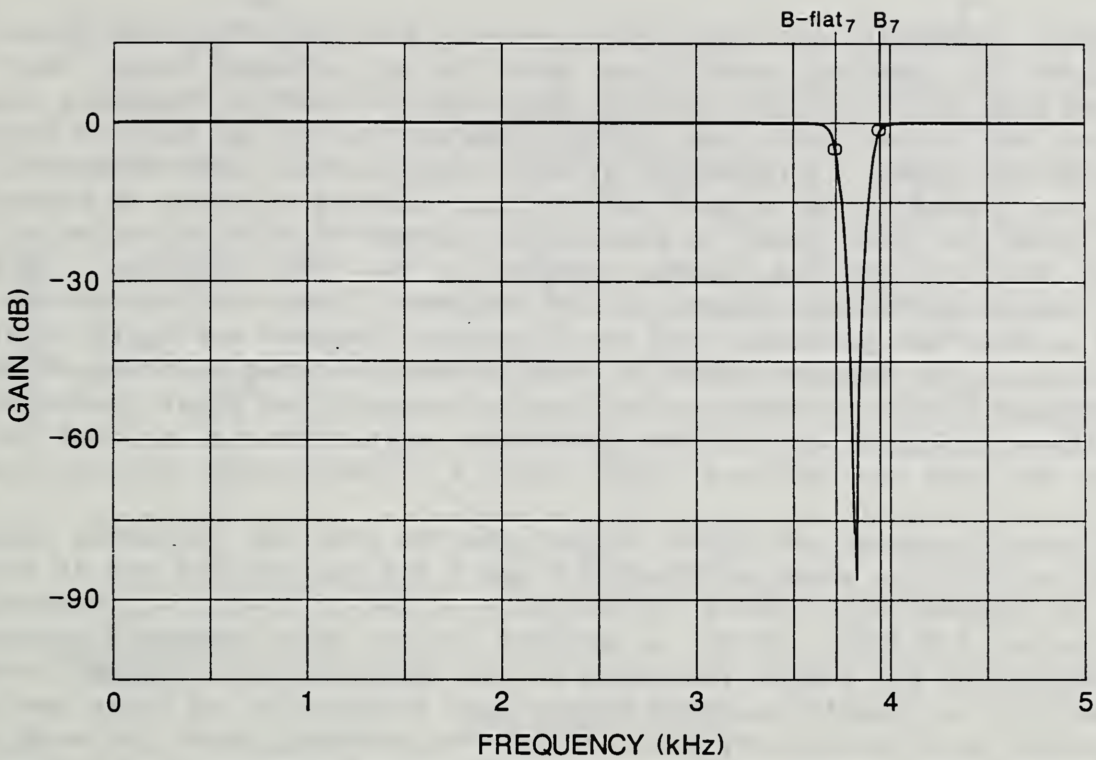


Figure 2 Amplitude-frequency characteristic for the encoding notch filter in the left channel of Encoder 006.

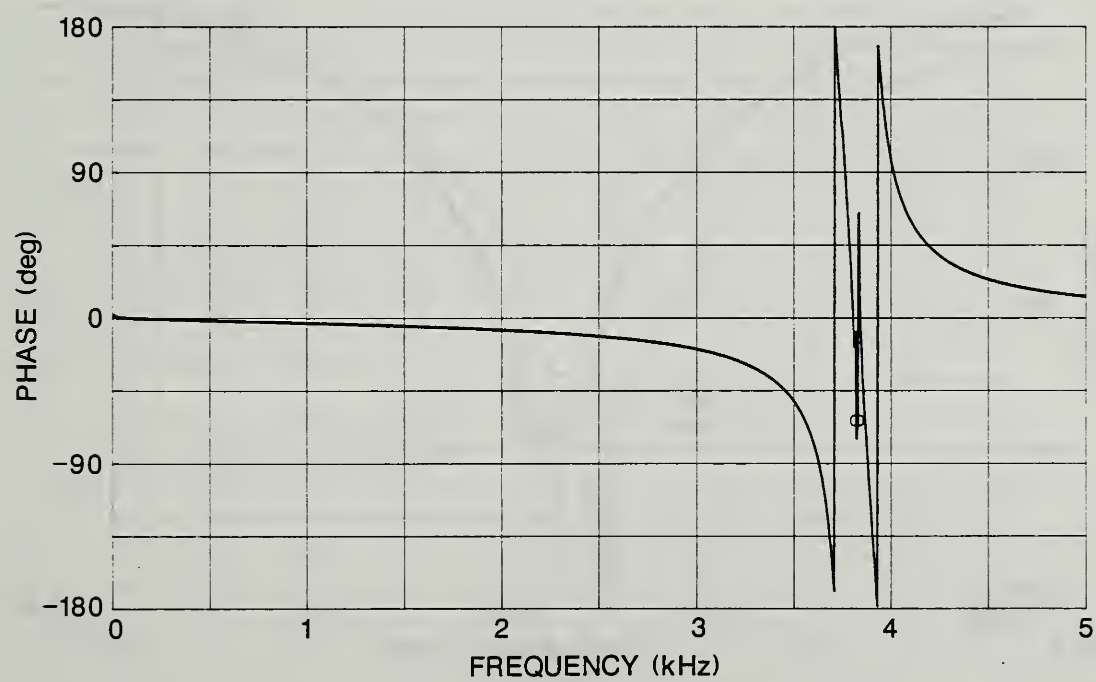


Figure 3 Phase characteristic for the encoding notch filter in the left channel of Encoder 006. (The small circle represents the cursor of the analyzer being at 3.84 kHz.)

To determine accurately the amplitude-frequency and phase response characteristics of the encoding notch filter over its full dynamic range, the swept-sine mode of the signal analyzer was used. A limited frequency range of 3690 Hz to 3990 Hz was used, and a pilot tone at 2740 Hz was used to force encoding at all times. A rejection filter, tuned to that same frequency of 2740 Hz, was placed in the signal path to the analyzer in order to prevent overload from the pilot tone. A frequency response of this rejection filter was first obtained with the Encoder switched to the "OUT" position. This response was stored in the memory of the analyzer. Then the Encoder was switched to the "IN" position, and the frequency response was again obtained. The division of the complex values of this second response by those of the first response yields the desired amplitude-frequency and phase characteristics of the encoding filter. These procedures were repeated, so that the responses for both the left and right channels of the Encoder were obtained.

The amplitude-frequency and phase characteristics over the frequency region from 3690 to 3990 are shown in figures 4 and 5 for Encoder 006 and in figures 6 and 7 for Encoder 012. These figures have marked on them the frequencies for the notes of B-flat$_7$ and B$_7$, as defined for the equal tempered chromatic scale. Note that the center frequency of the encoding notch for each channel of Encoder 012 is closely centered between the frequencies of these two notes and each note is attenuated by about 3 dB. The encoding notch for each channel of Encoder 006 is centered at a somewhat lower frequency, and attenuates B-flat$_7$ by about 5 dB, and B$_7$ by about 2 dB. For both channels of both

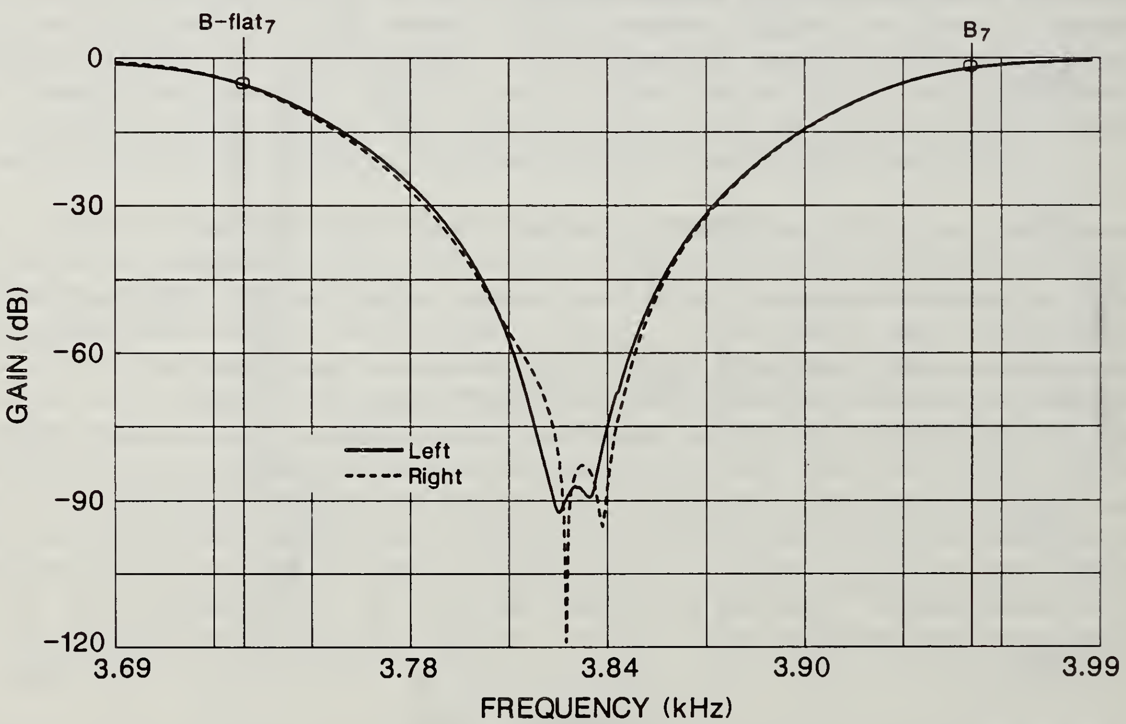


Figure 4 Amplitude-frequency characteristics of the encoding notch filters in Encoder 006. The solid and dashed curves correspond to the left and right channels, respectively.

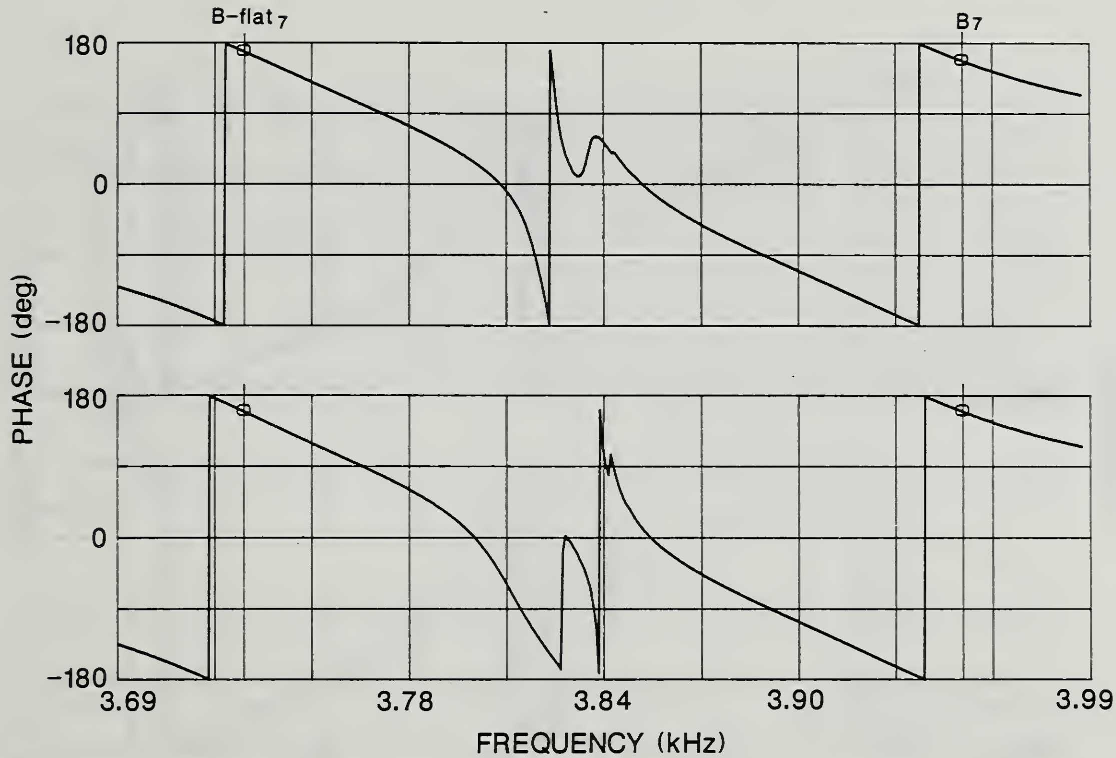


Figure 5 Phase characteristics of the encoding notch filters in Encoder 006. The upper and lower plots correspond to the left and right channels, respectively.

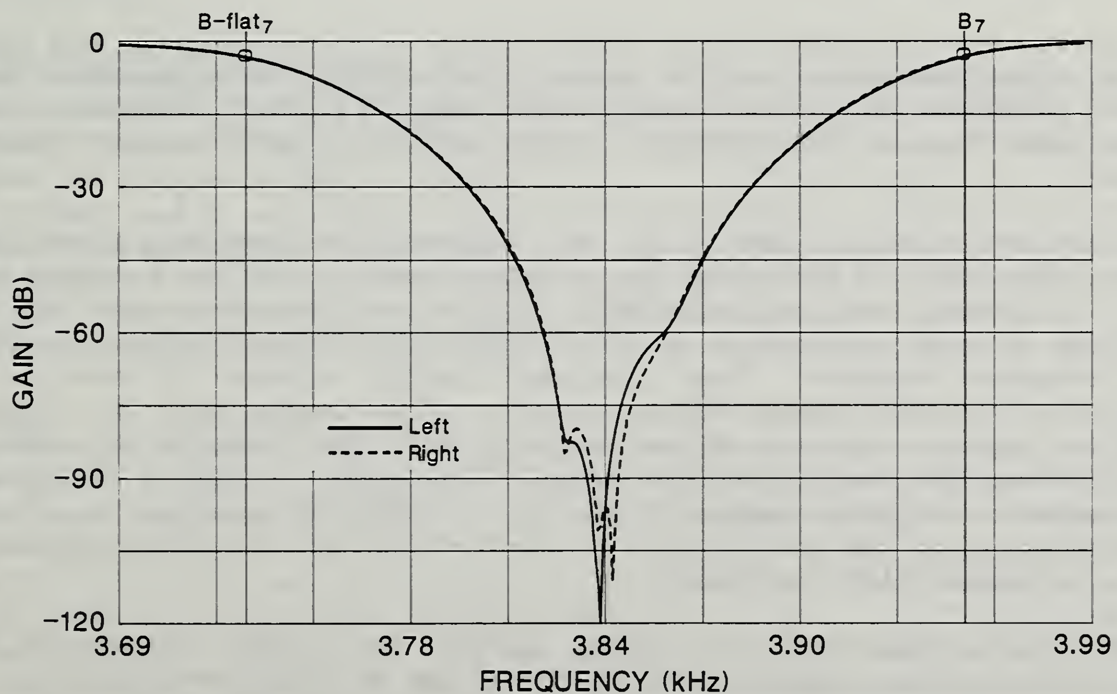


Figure 6 Amplitude-frequency characteristics of the encoding notch filters in Encoder 012. The solid and dashed curves correspond to the left and right channels, respectively.

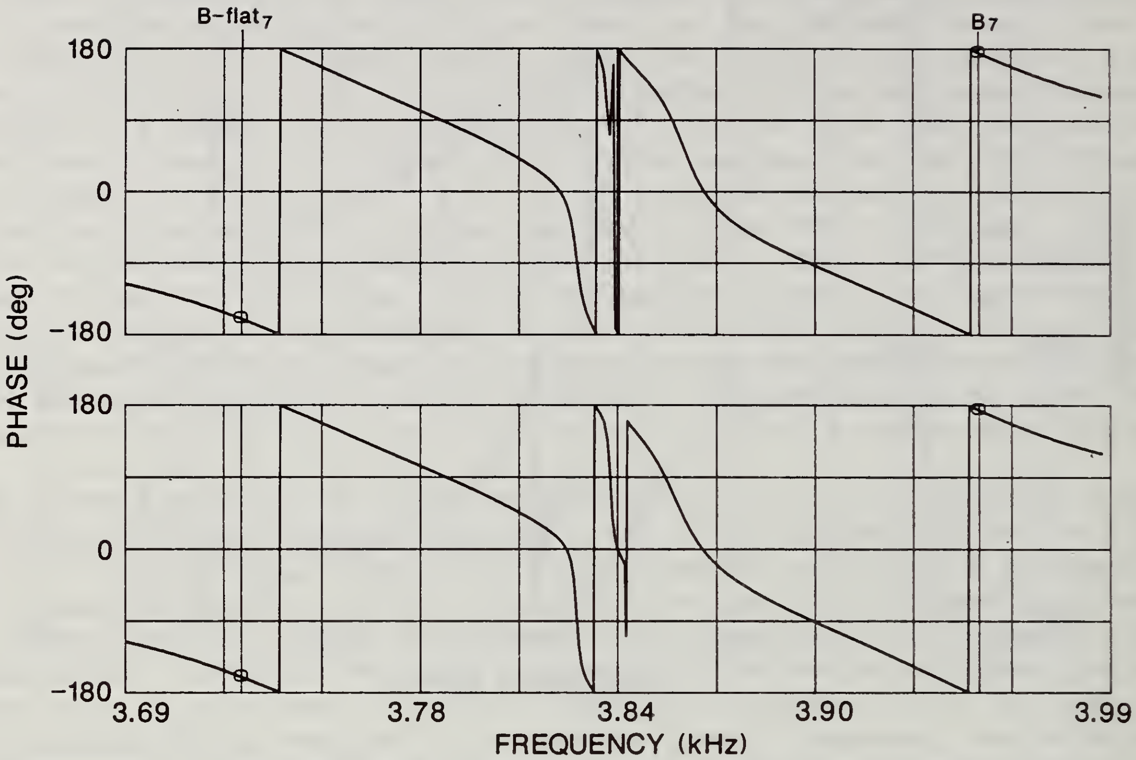


Figure 7 Phase characteristics of the encoding notch filters in Encoder 012. The upper and lower plots correspond to the left and right channels, respectively.

Encoders, the attenuation near the center of the notch-filter passband is in excess of 80 dB and the phase characteristic deviates drastically from a flat, uniform phase response, exhibiting a discontinuity in the frequency region of the notch.

Since a narrowband filter can "ring" when excited by a transient signal, measurements were made to determine the transient response of the encoding notch filter. The signal analyzer employed for these tests has the capability of determining an impulse response by calculating the inverse Fourier transform of the frequency response. This calculated impulse response is shown in figure 8 for the left channel of Encoder 006. The ringing which is seen after the initial impulse consists of two frequencies in the vicinity of 4 kHz. The beating between the two frequencies causes the ringing to appear to decrease and increase in amplitude several times before dying out completely. The impulse response of the left channel of Encoder 012 was also measured and found to be essentially identical.

In order to be certain that the ringing observed was not an artifact of the filters in the signal analyzer rather than of the encoding notch filter, direct time-domain measurements were also made. The ringing behavior observed in these measurements was similar to that shown in figure 8.

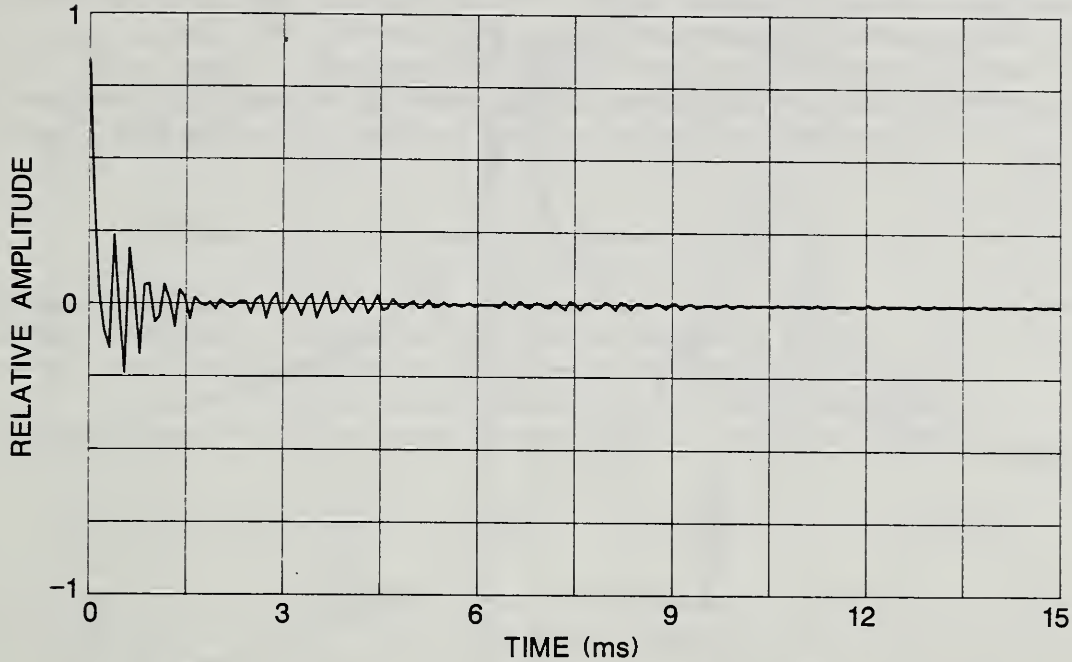


Figure 8 Impulse response, on a linear scale, of the left-channel notch filter for Encoder 006 as obtained from the inverse Fourier transform of the frequency response.

<u>Logic circuitry</u>

<u>On-notch and off-notch filters</u> These filters (see figure 1) were characterized by placing a 1-megohm probe at their outputs and obtaining their frequency responses using the swept-sine mode of the signal analyzer. The amplitude-frequency characteristics of these filters for Encoders 006 and 012 are shown in figures 9 and 10, respectively. The frequencies for the musical notes B-flat$_7$ and B$_7$ are shown for the on-notch filter. The frequencies for the musical notes E$_7$ and F$_7$ are shown for the off-notch filter. Notice that the maximum level from the on-notch filter of Encoder 012 is several decibels lower than that of Encoder 006. This reduced output caused Encoder 012 to keep the encoding filter engaged more of the time than does Encoder 006.

<u>Switching</u> The signal levels that cause the encoding filter to be switched in or out were determined using two sine-wave generators, one adjusted to produce a tone at the frequency of maximum response of the on-notch filter of the Encoder under test and the other set to the frequency of maximum response for the off-notch filter. The summed output of these frequency synthesizers was fed to both the left and right inputs of the Encoder. The input level of the on-notch tone was adjusted to a particular value, e.g., -3 dBV (where here and elsewhere in this report, "dBV" denotes an rms-voltage level, relative to 1 V rms, expressed in decibels). The level of the off-notch tone was set at that same value and then slowly reduced in 1 dB increments until encoding ceased. The level of the off-notch tone was then reduced several decibels further, and then slowly increased until encoding resumed. A new level was selected for the on-notch frequency, and the process was repeated.

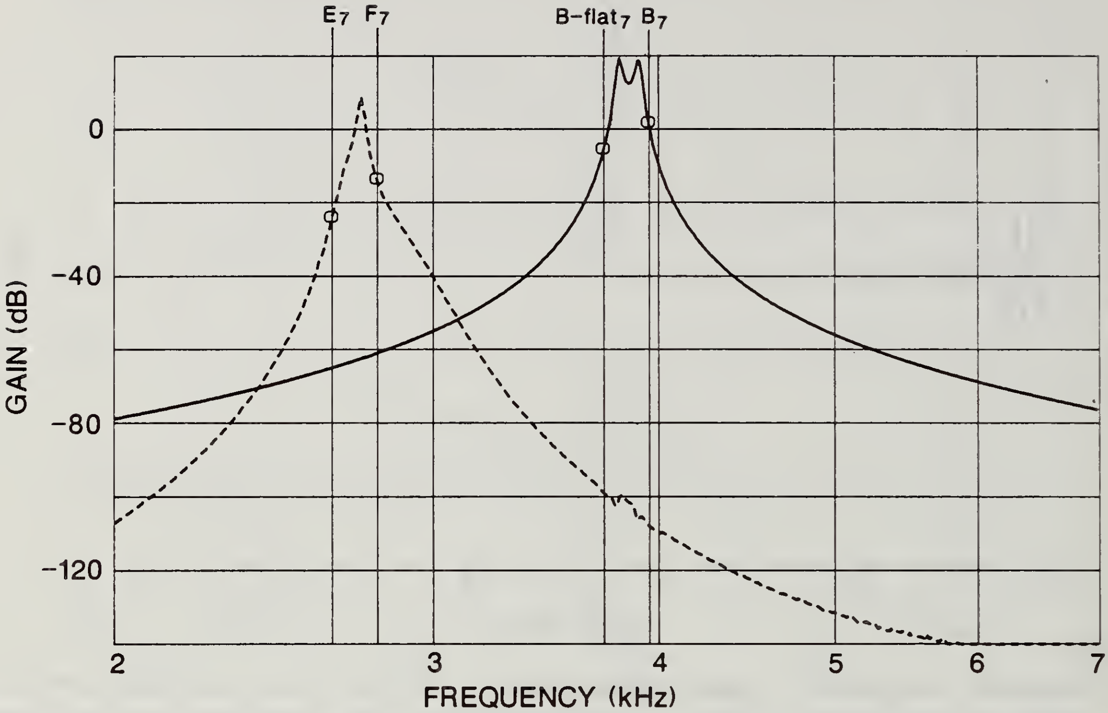


Figure 9 Amplitude-frequency characteristic of the on-notch (solid curve) and off-notch (dashed curve) filters in Encoder 006.

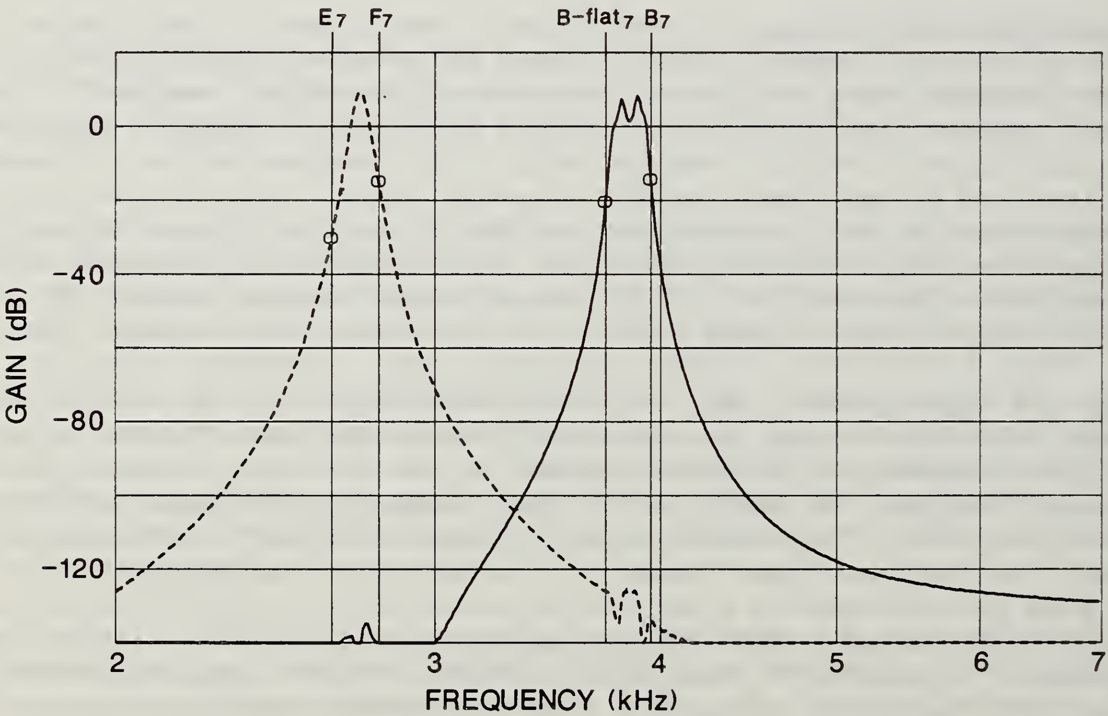


Figure 10 Amplitude-frequency characteristic of the on-notch (solid curve) and off-notch (dashed curve) filters in Encoder 012.

Tables 1 and 2 show the results of these measurements of the switching action for Encoder 006 and 012, respectively. These values were not completely stable, particularly for Encoder 006 where the results varied with time over a range of 1 dB or more. The reduced output from the on-notch filter of Encoder 012 as compared to that of Encoder 006 is reflected in the difference between the on-notch and off-notch action levels in the two encoders.

Table 1. Logic map for Encoder 006

On-notch voltage level at input	Off-notch voltage level at input required to change encoding state	
	to start encoding	to stop encoding
-3	-4	-8
-13	-15	-19
-23	-25	-29
-33	-32	-41
-38	-42	-52
-40	Always encodes	

Table 2. Logic map for Encoder 012

On-notch voltage level at input	Off-notch voltage level at input required to change encoding state	
	to start encoding	to stop encoding
7	2	-6
-3	-13	-16
-13	-23	-26
-24	-36	-40
-27	Always encodes	

Note:the use of the term "dBm" has been avoided in these tables since the quantity measured was voltage, not electrical power. Since the term "dBm" has been used in the audio industry to denote voltage level referenced to 1 milliwatt of average electrical power dissipated in a purely resistive 600-Ω electrical load, the following is included for the reader's information. The rms voltage corresponding to 1 milliwatt of average electrical power dissipated in a purely resistive 600-ohm electrical load is approximately 0.775 V rms or, in logarithmic terms, -2.2 dBV.

During the course of these measurements of the switching action, it was noted that the encoding notch filters were switched in and out over a period of some milliseconds so that abrupt switching transients did not occur.

Figure 11 shows the voltage levels at the outputs of the on-notch and off-notch filters when a 20-second sample of Prokofiev's "Alexander Nevsky" (see Appendix E for a description of music used in this study) was played through Encoder 012. The presence of an orange line at the bottom of the figure indicates that the encoding notch filter was engaged; when no orange line is present, the Encoder had dynamically removed the notch filter from the signal path. Figure 12 is a similar plot for a sample of glockenspiel music passed through Encoder 006, showing a change from continuous encoding to intermittent encoding at a time when there was an obvious change in the musical key. As can be seen, each of the Encoders could change states frequently within a short period of time. Because of the higher voltage levels from the on-notch filter in Encoder 006, this unit frequently would switch the encoding notch filter out under conditions when Encoder 012 would continue to encode.

It was noted that the Encoder action was not always repeatable; immediate repetition of a sample of music did not always lead to the Encoder changing state at the same time in the music.

Dynamic range, Frequency Response, and Distortion

With the encoding switch in the "OUT" position so that the left- and right-channel signals passed through the unity gain amplifiers of the Encoder but the encoding notch filter was not engaged, the dynamic range of the Encoder was measured and found to be in excess of 105 dB.

The frequency characteristic of each Encoder was measured with the the encoding switch in the "OUT" position. Figures 13 and 14 show the measured amplitude-frequency characteristic and the phase characteristic, respectively, for the left channel of Encoder 006, as determined over the frequency range from 125 Hz to 100 kHz using the HP 3562A analyzer with its random noise source. The root-mean-square input voltage for this measurement was 0.77 V. The shape of the response may be taken as typical of both channels of both Encoders. Switching the Encoder "IN" affects the response only in the region near 3.84 kHz, as indicated in figures 2 through 7.

Harmonic distortion of the Encoders, with the encoding switch in the "OUT" position, was checked by using an input signal from a low-distortion oscillator and the HP 3562A analyzer in the THD mode. The results, for all combinations of frequency and levels which were tried, indicated that total harmonic distortion was more than 80 dB below the level of the fundamental frequency.

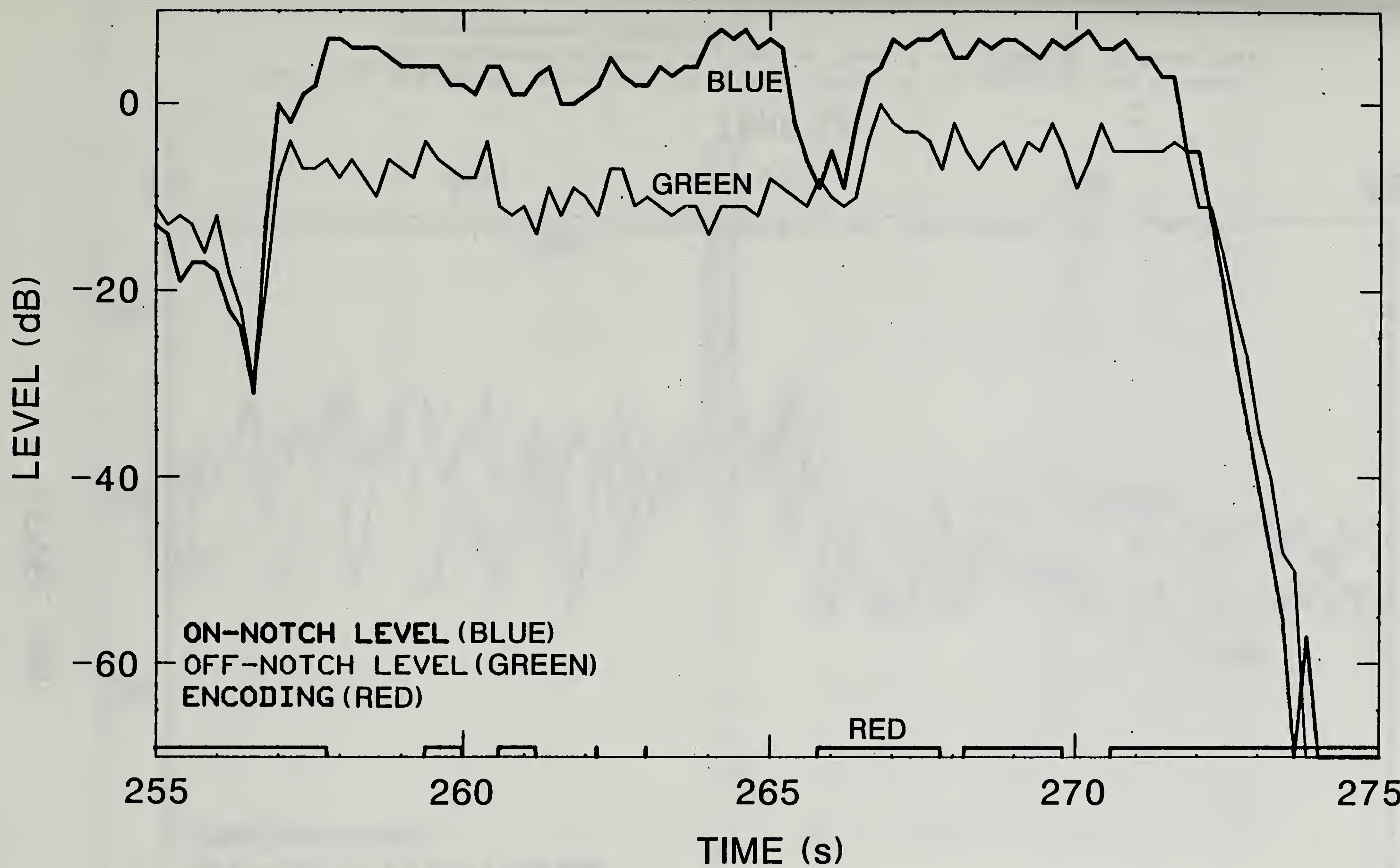


Figure 11. The Encoder state and the voltage levels at the output of the on-notch and off-notch filters for Prokofiev selection (see Appendix E for a more complete description of this selection).

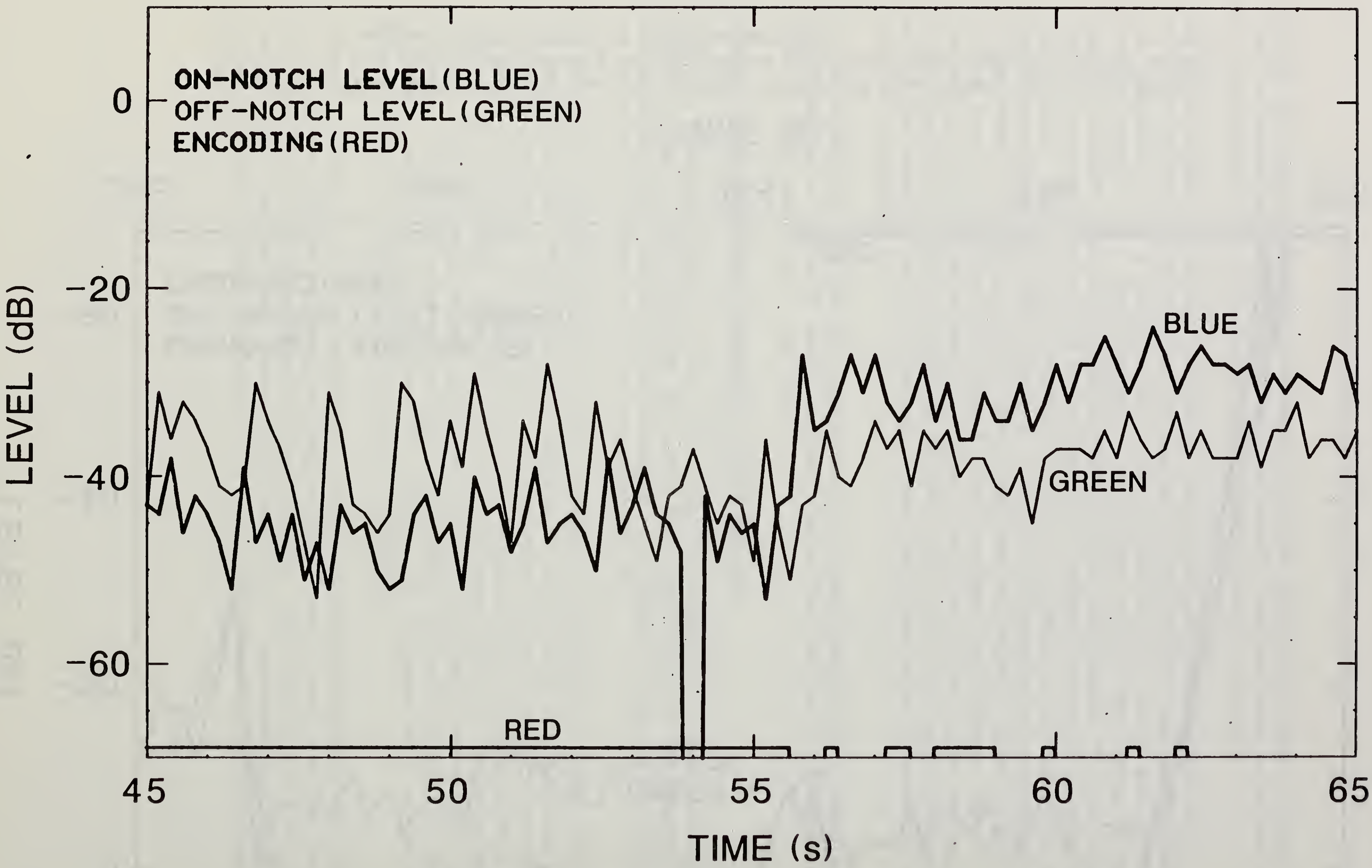


Figure 12. The Encoder state and the voltage levels at the output of the on-notch and off-notch filters for a part of Track 3 of "Drumming" by Steve Reich (Electra/Nonesuch 9079170-2).

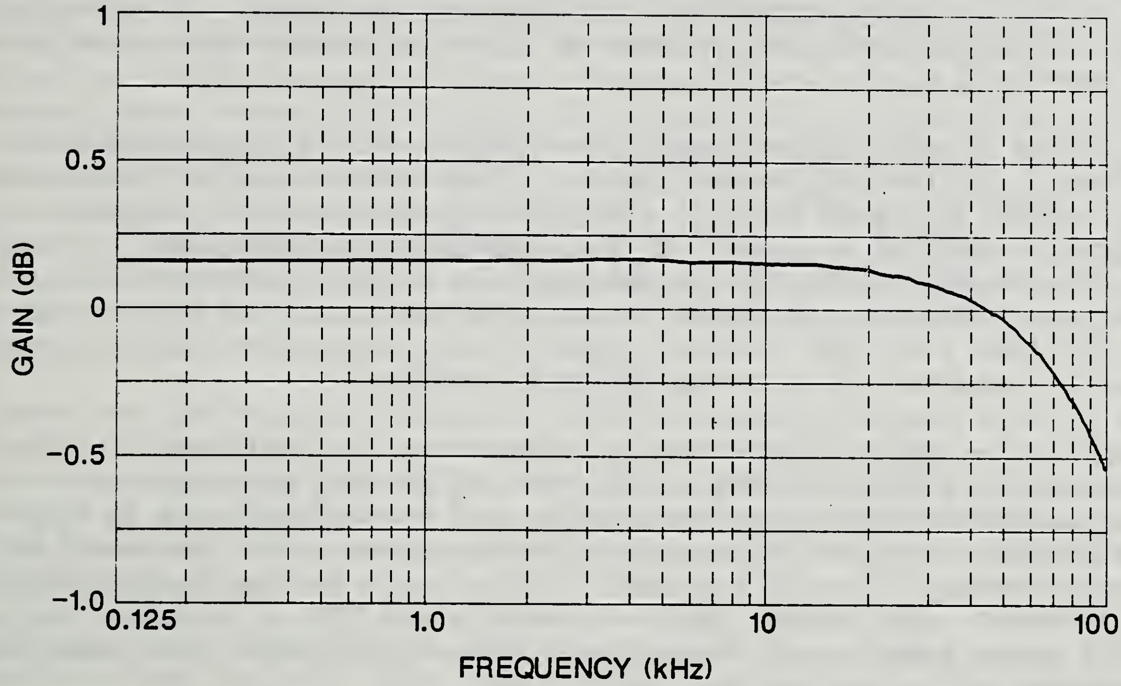


Figure 13 Amplitude-frequency characteristic for the left channel of Encoder 006 with the encoding notch filter out of the circuit.

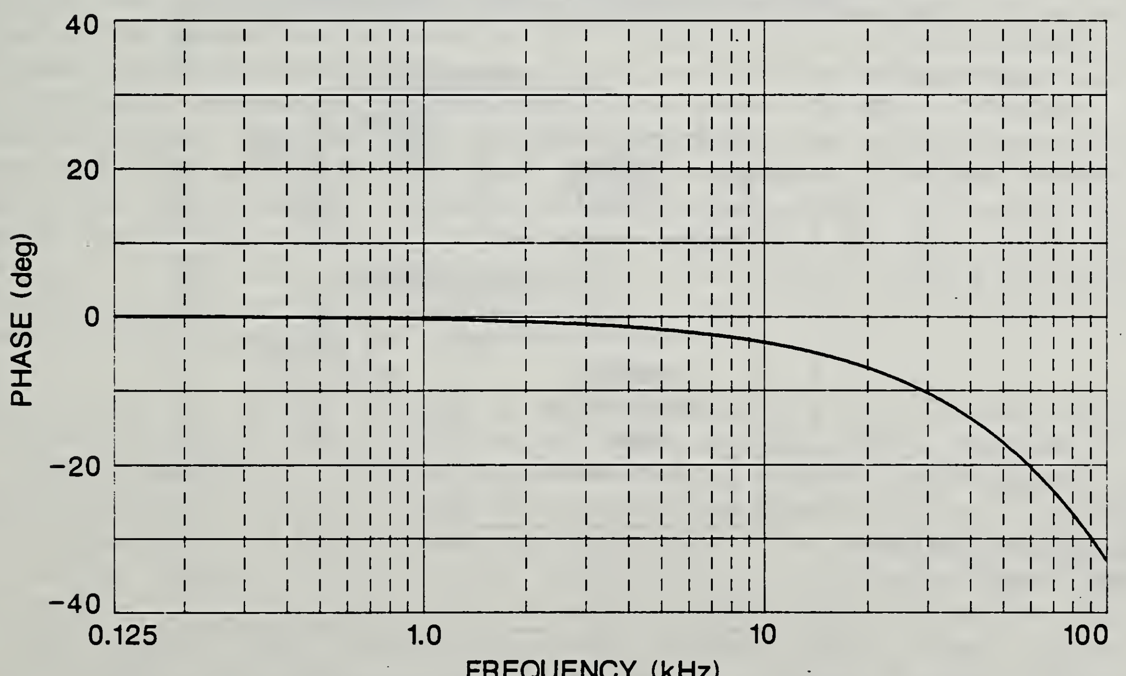


Figure 14 Phase characteristic for the left channel of Encoder 006 with the encoding notch filter out of the circuit.

2.3 Decoder: Operation

A general description of the operation of the CBS-supplied DAT Recorder, modified to contain a "copy prevention code" scanner or Decoder, is given in this section. Details about the performance of the DAT Recorder/Decoder are provided in section 2.4.

The purpose of the Decoder is to detect the presence of the copy prevention code contained in the incoming audio signals. This information or code is the removal of a narrow frequency band of signal amplitudes from the original, unencoded signal, creating a "notch" in the frequency spectrum. As implemented for the NBS tests, the Decoder portion of the DAT Recorder is simply an adjunct circuit to the other recorder electronics, as illustrated in figure 15. The left and right channel signal lines coming into the recorder are connected in parallel to go to the Decoder circuits.

If the presence of a notch is detected by the Decoder, it activates a yellow light-emitting diode (LED), mounted on the rear of the DAT Recorder, to indicate the possibility of the presence of a copy prevention notch in the signal. The Decoder then continues to monitor the presence of the notch for a period of approximately 13 to 15 seconds. If the notch in the signal persists during this interval, the Decoder then activates a red LED to indicate that a RECORD INHIBIT state exists, and this signal is used to disable the other DAT Recorder electronics.

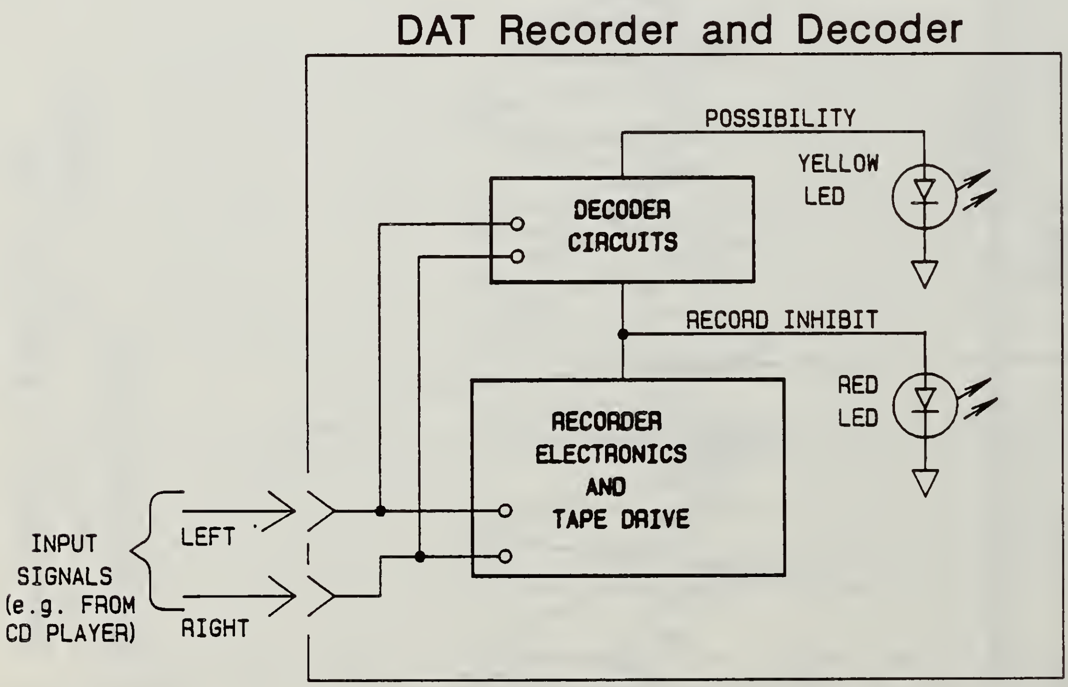


Figure 15 Basic diagram of the decoder and recorder electronics in the CBS-supplied DAT Recorder/Decoder.

The Decoder operates by passing the audio input signal through three bandpass filters, creating three separate signals having levels proportional to the amplitude of the input signal frequencies contained in the narrow passband of each filter. The filters are tuned to the notch center frequency and to frequencies slightly above and below the notch center frequency. The frequency bands above and below the notch band are referred to as the upper and lower sidebands, respectively. By comparing signal levels corresponding to the upper and lower sideband filters with the signal corresponding to the notch bandpass filter, the Decoder detects when the level in the notch frequency passband is less than the level in either of the two sidebands, indicative of a notch present in the frequency spectrum of the incoming audio signal. A more detailed description of the operation of the Decoder is given below.

In order for the Decoder to operate when the incoming signal has its frequency spectrum shifted (due to variations in the encoding process or the original source material), the Decoder has a built-in scanner which provides for the detection of a notch which is 5 to 10 percent above or below the nominal center frequency. The scanning time to cover this variation was found to be approximately 35 seconds.

The Decoder must differentiate between a naturally-occurring notch in the spectrum of the original audio input signal and one due to the encoding process. To make this distinction, NBS found that the Decoder implements several criteria. One requirement is the time during which the notch must be sustained. As mentioned above, this time is about 13 to 15 seconds. Another criterion is a minimum level of the audio input signal in the upper and lower sidebands just above and below the notch. The level of the sum of the signals in these two sidebands must be above a lower threshold limit in order to activate the comparison process in the Decoder, i.e., a notch cannot be detected simply because the input signal level in the notch region is too low. This low-level threshold was measured to be about -51 dBV. Finally, as described above, the level of the audio signal in the notch frequency band is compared to the level of the signal in the sidebands. For the Decoder to detect (and indicate) the presence of an encoded signal, the notch band level must be lower than the level of the lower sideband. As will be described in section 2.4, physical measurements made on the Decoder showed considerable variation in the required relative levels of the notch and sideband signals in order for this criterion to be met.

Figure 16 on the following page is a block diagram of the Decoder circuitry contained in the DAT Recorder/Decoder. The incoming audio signals (both left and right channels of a stereo input) are first summed together and then passed through a low-pass, anti-aliasing filter which produces a combined audio signal free of high-frequency components. The combined signal is then fed into an automatic gain control (AGC) circuit. The function of the AGC circuit is to keep the level of its output signal within a narrow range, while its input signal level varies over a wide dynamic range. Next, the AGC output signal is modulated using a carrier frequency of about 10 kHz. This modulation shifts the frequency spectrum of the combined audio signal to frequency bands above and below the modulation frequency. The analog

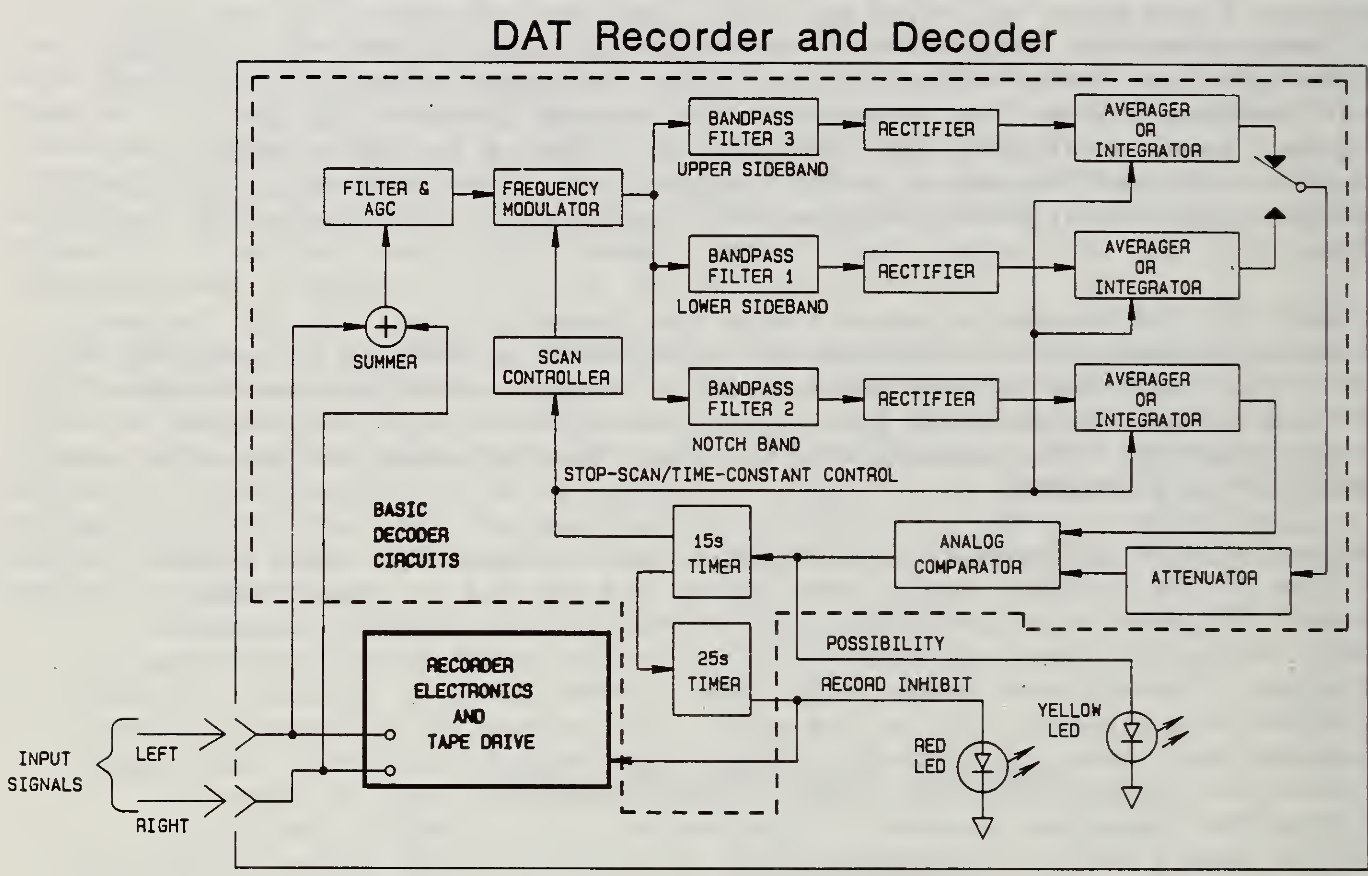


Figure 16 Block diagram of the major circuits contained in the Decoder section (within the dashed outline) of the DAT Recorder/Decoder, as determined in the NBS study. These circuits were implemented on an ancillary circuit board, separate from the other recorder electronics and tape drive.

discrimination circuits following the modulator are tuned to make use of the audio signal spectrum that lies below the modulation frequency. The frequency-shifted signal is then split into three signal paths, as indicated in figure 16, and fed into three separate bandpass filters. Each filter is tuned to one of three center frequencies, corresponding to the upper sideband, lower sideband, and notch band. The modulation frequency is slowly varied by the scan controller circuit so that the bandpass filters sample a varying frequency range of the incoming audio signal.

Following the bandpass filters, the resultant narrowband signals, centered about their corresponding bandpass filter center frequency, are then converted to dc signals (by a rectifier circuit and low-pass filter). Thus, the audio input signal is decomposed into three dc signal levels, each proportional to

the amplitude of the input signal frequencies contained in the particular passband (notch band, upper and lower sidebands). These dc signals are next processed by either averaging or integrating the three signals, depending on the state of the stop-scan/time-constant control line, i.e., whether or not a notch in the input signal has been detected by the subsequent analog comparator and digital logic circuits. If a notch has not been detected, the Decoder is in the scanning or "search" mode. If a notch has been detected, the Decoder is in the integrate mode.

In the scanning mode, the three dc signals are averaged with circuits having short time constants (100 to 300 ms). The smaller of the upper or lower sideband signals is switch selected by another sideband selector comparator circuit, not shown in this diagram. The upper or lower sideband averaged signal is first attenuated, and then compared by the analog comparator with the averaged notch signal. The attenuator essentially provides the difference factor (or offset) by which the averaged notch signal must be less than the averaged sideband signal in order for the comparator to indicate the presence of a notch in the spectrum of the input signal. When the level of the averaged dc signal from the notch band filter falls sufficiently below the level of the averaged dc signal of the selected sideband, the comparator output line changes, activating the rear panel yellow LED, and indicating the "POSSIBILITY" state. The change in output from the analog comparator also causes the 15-second timer to start its timeout cycle.

One of the outputs from the 15-second timer is the stop-scan/time-constant control line so that upon changing its state, the scan control circuit stops the frequency sweep, and the averaging circuits are switched to true integrators. The Decoder is then in the integrate mode (usually referred to as the POSSIBILITY state). If the input signal conditions persist such that the comparator continues to detect the presence of a notch, while the 15-second timer completes its full 15 second timeout cycle, then another output from the 15-second timer activates the 25-second timer. As soon as it is activated, this timer changes its output to provide the "RECORD INHIBIT" state, which activates the rear panel red LED indicator and causes the recorder electronics to shut down for a minimum of 25 seconds. At the end of that time, if the comparator continues to detect the presence of a notch, another 25 second shutdown begins immediately.

2.4 Decoder: Physical Measurements

Frequency Response of the Three Bandpass Filters

The amplitude-frequency characteristics of the three bandpass filters used to decompose the output signal from the frequency modulator into three separate narrowband signals, as described in section 2.3, is critical to the performance of the Decoder. It is the shape of these filters, and their relative position on the frequency scale, that determines the spectrum and amount of energy from the input signal that is required to cause the POSSIBILITY and RECORD INHIBIT states of the Decoder to occur.

Figure 17 is a plot of the amplitude response of the three bandpass filters in the Decoder versus the effective input signal frequency. These characteristics were measured by a spectrum analyzer with the recorder output from the spectrum analyzer used to generate the corresponding plots. The notch bandpass filter is indicated in green, the upper sideband filter in blue, and lower sideband filter in red. These filters are tuned to capture signal frequencies that are below the modulation frequency, which was measured and found to vary from 9.9 to 10.5 kHz. The corresponding input signal frequency has been derived from the measured filter frequency by subtracting the measured filter frequency from the nominal 10.2 kHz modulation frequency.

These are high Q filters ($Q = \sim 80$) with 3 dB bandwidths of about 50 Hz. The skirts of these filters begin to overlap at rejection levels that are 30 dB or more down from the peaks. The non-ideal ripple in the skirts below these levels has negligible effect on the Decoder performance. The three bandpass filters of the Decoder have excellent narrowband rejection characteristics that select distinctive frequency segments of the incoming signal to be selected for purposes of notch encoding discrimination. Note that the peak of the upper sideband filter is a few dB lower than that of the notch and lower sideband filters. This attenuation appears to compensate for the gain emphasis of the input filter and AGC characteristic in the frequency region of the upper sideband, as described below.

Frequency Response of the Anti-Aliasing Filter and AGC Circuit

As described in section 2.3, the operation of the Decoder includes the use of an anti-aliasing filter, following the input summing circuit, to produce a combined audio signal free of high frequency components. Measurements on the automatic gain control (AGC) circuits following the filter showed that the AGC output signal amplitude varied linearly with the AGC input signal amplitude. That is, the AGC circuit of the Decoder was found to be inoperative (examination with the aid of the circuit schematic provided by CBS showed that a connection had been made that disabled the AGC circuits).

Figure 18 shows the plot of the amplitude-frequency characteristic of the anti-aliasing filter and AGC circuits of the Decoder. In the notch region from 3 kHz through 5 kHz the response is nearly flat, except around 4.5 kHz where there is several dB of relative gain emphasis. Above 5 kHz the response rolls off smoothly to reject high frequency components in the audio input signal. Similarly, below 3 kHz the low frequency components of the input

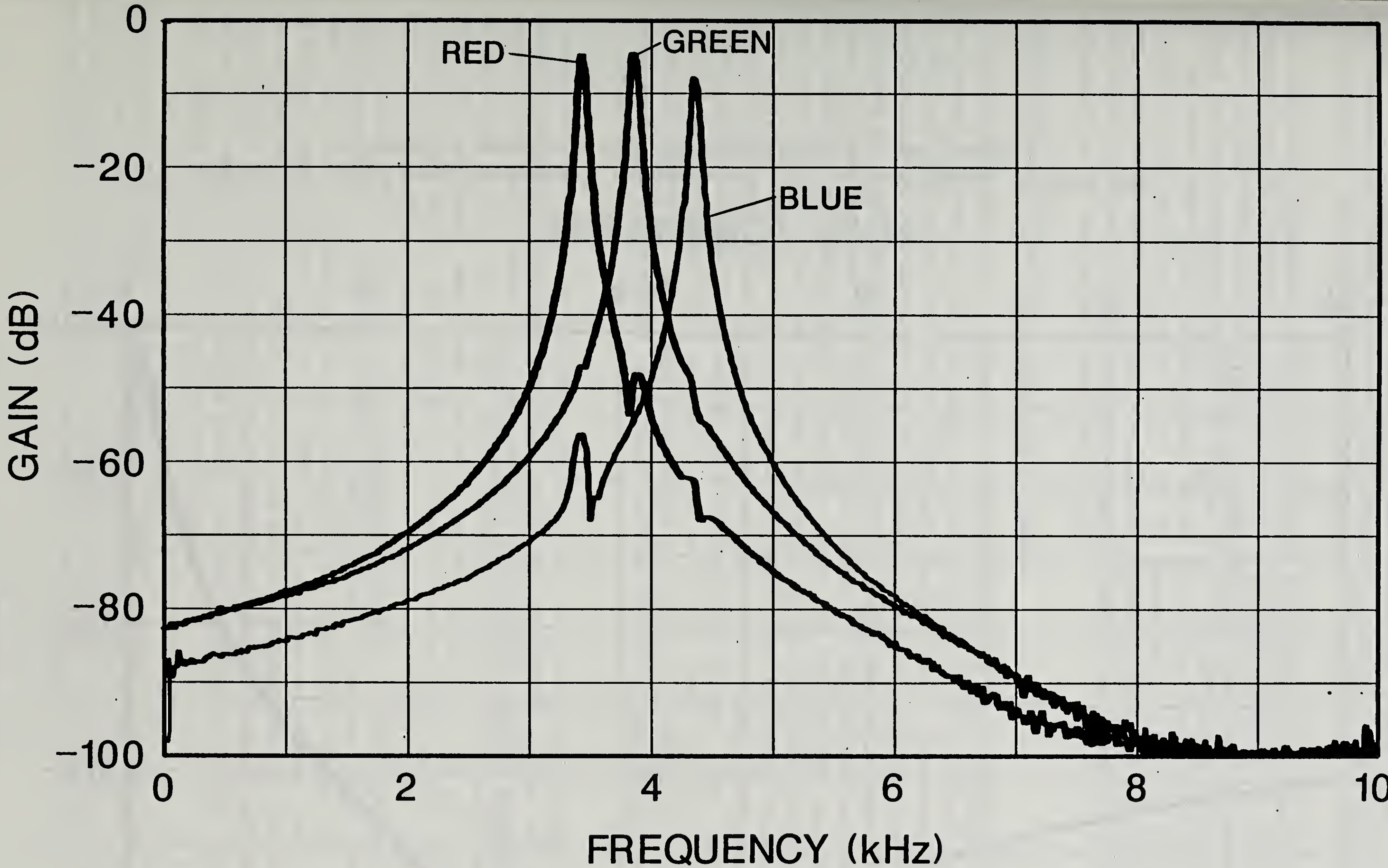


Figure 17 Amplitude response for the three bandpass filters in the Decoder versus the effective input signal frequency.

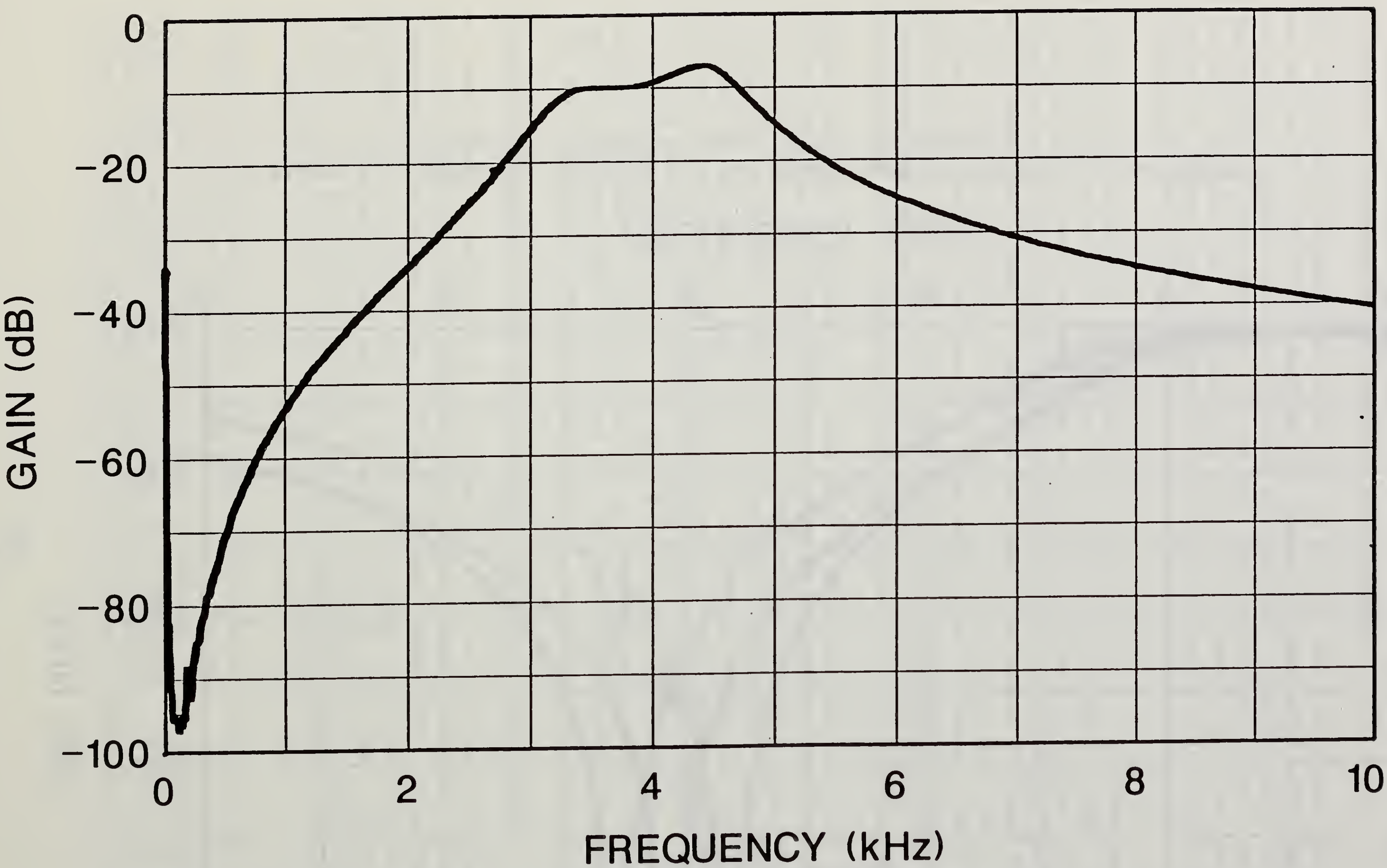


Figure 18 Amplitude-frequency response of the anti-aliasing filter and automatic gain control (AGC) circuits in the Decoder.

signal are rejected, particularly those around the power line frequency, so as to prevent any unwanted 60 Hz (and related harmonics) from affecting the operation of the discrimination circuits.

By comparing figures 17 and 18, the total frequency characteristics of the Decoder can be inferred. The total input signal spectrum is first attenuated by the characteristic of figure 18, and then the signal is separated into the three narrow frequency bands of figure 17. The frequency sweeping operation of the scan controller effectively moves the three bandpass filter characteristics slowly over about 300 Hz above and below the nominal frequency position shown here. The gain emphasis of the filter/AGC characteristic in the upper sideband region is thus compensated by the slightly greater attenuation of the upper sideband filter.

Steady-State, Three-Tone, Sine-Wave Signals

As described above in section 2.3, the Decoder circuit responds to the frequency components of the input signal primarily in the vicinity of the notch band and upper and lower sideband frequencies. Consequently, the performance of the Decoder can be characterized on a steady-state basis by applying an input signal composed of three superimposed sinusoidal waveforms whose frequencies are at (or near) the center of the three bandpass filters of the Decoder, and whose amplitudes can cause the POSSIBILITY and RECORD INHIBIT states to occur. For the purpose of these tests the scan controller was deactivated. The three-tone, sine-wave signal had its three frequencies adjusted to correspond to the center of the three bandpass filter characteristics that were measured and described above. The level of the notch frequency input was varied relative to the level of the upper and lower sideband inputs to determine the levels at which the POSSIBILITY state was activated and inactivated, and the level below which the RECORD INHIBIT state occurred.

Figure 19 is a plot of the data obtained on the level differences between the notch frequency input and the smaller of the sideband inputs for which the POSSIBILITY (activated and inactivated) and RECORD INHIBIT states were observed as a function of the lower sideband level. For these measurements the upper sideband amplitude was held at approximately -43 dBV. The level difference at which the POSSIBILITY state is activated (green curve) is about 2 to 4 dB below the level difference at which the POSSIBILITY state is inactivated (blue curve). More important, however, is the plot of the level difference that causes the RECORD INHIBIT state to occur (red curve). Note how this curve deviates from the POSSIBILITY activated state for lower sideband levels above about -53 dBV. As the lower sideband level increases, the difference level required to cause RECORD INHIBIT decreases from the POSSIBILITY activated level up to a sideband level of about -23 dBV where it drops off rapidly until, above about -16 dBV, a RECORD INHIBIT state could not be obtained at any (low) difference level.

Analysis of the block diagram of the Decoder circuitry given in figure 16 indicates that the difference characteristic given in figure 19 should depend essentially on the operation of the attenuator and analog comparator circuits. As described in section 2.3, the attenuator provides the difference factor (or

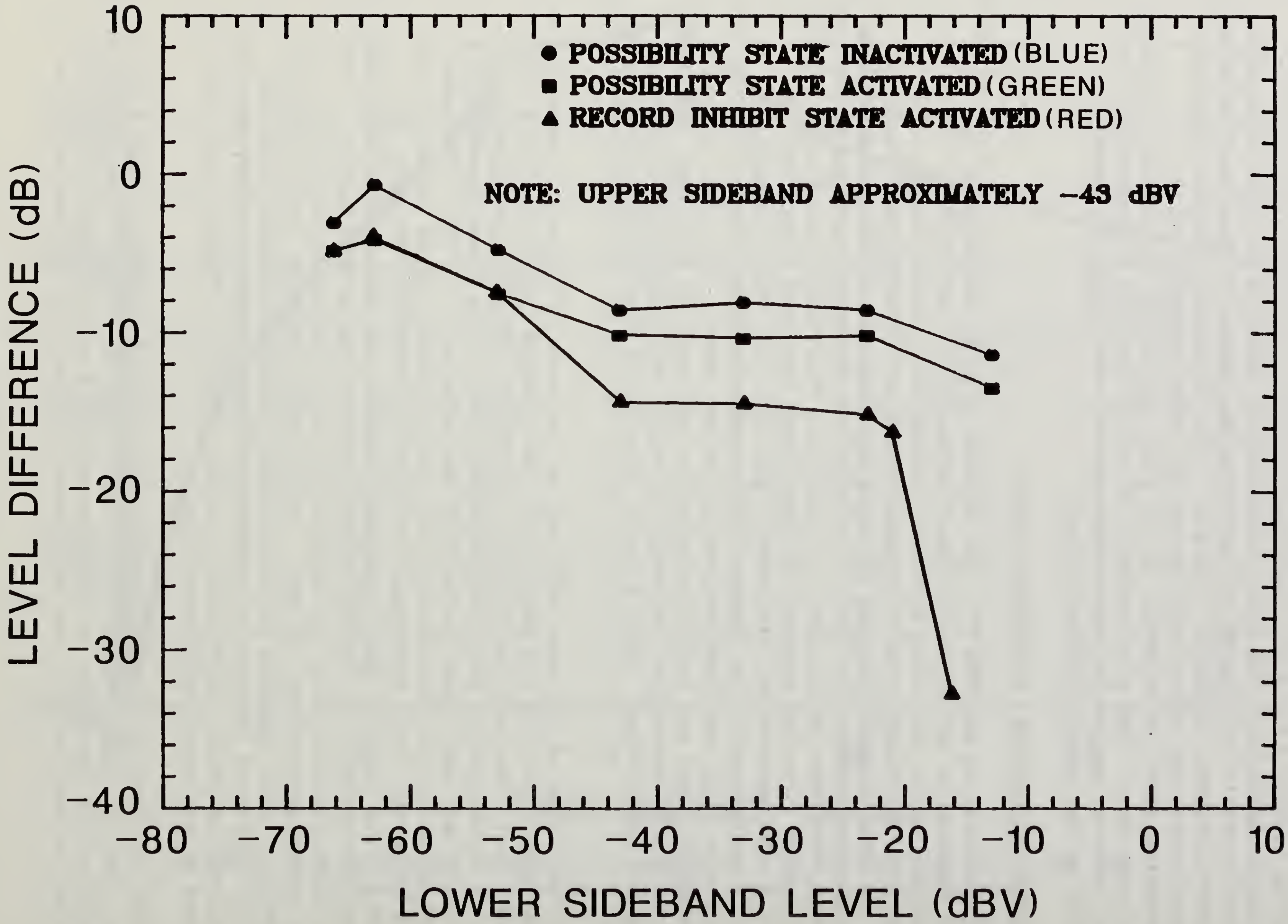


Figure 19 Difference level as a function of the lower sideband level (upper sideband level constant at about -43 dBV) that causes the POSSIBILITY AND RECORD INHIBIT states in the Decoder.

offset) that the analog comparator uses to make its threshold level comparison. On this basis, this difference level comparison would be expected nominally to be constant over the operating range of input levels, which would yield a flat line difference characteristic with the POSSIBILITY activated and RECORD INHIBIT states at the same level. NBS found that the actual difference characteristic exhibited by the Decoder (as shown in figure 19) is more complex, in part due to insufficient dynamic range for making the difference level comparison (as will be described), and to the limitations of actual (nonideal) circuit components.

During the POSSIBILITY state the output voltage level of the integrators slowly increases. With the lower sideband input level above about -53 dBV, it was observed that the output of the lower sideband integrator saturates at ~14 V dc before the 15 second timeout, with a corresponding output from the attenuator of ~3.4 V dc. Since the output voltage of the notch band integrator does not get attenuated, it also saturates at ~14 V dc. Therefore, whenever the notch band integrator output voltage exceeds ~3.4 V dc, the analog comparator output changes out of the POSSIBILITY state (integrate mode) into the scanning mode. If this change occurs before the 15 second timer completes its timeout cycle, the RECORD INHIBIT state is not reached. With a lower sideband input level above about -53 dBV, the ~3.4 V dc maximum output level from the attenuator is exceeded before 15 seconds, so that if the output from the notch band integrator continues to increase above ~3.4 V dc, the analog comparator changes state. Therefore, the difference level between the notch signal input and the lower sideband input must be smaller than that required to cause the POSSIBILITY activated state, in order to prevent saturation and to cause a RECORD INHIBIT state.

Next, the same three-tone, sine-wave signal was applied to the Decoder and the notch amplitude varied as above, only with the upper sideband amplitude held at approximately -23 dBV. This characteristic is plotted in figure 20. It differs considerably from the characteristic described in the previous paragraphs. In this case, the level differences observed for obtaining the POSSIBILITY state are more variable, and at a somewhat lower level. The difference level required to cause RECORD INHIBIT also deviates considerably more from the POSSIBILITY activated level with the lower sideband level above -53 dBV. The RECORD INHIBIT state could not be obtained for any (low) difference level when the lower sideband level was greater than about -21 dBV.

Figures 21 and 22 correspond to figures 19 and 20, respectively, and are the plots obtained from the data taken on the level differences for the three Decoder states (POSSIBILITY activated, POSSIBILITY inactivated, and RECORD INHIBIT), as described above, except that the lower sideband is held constant at -43 and -23 dBV levels while the upper sideband level is varied. Very similar results were obtained for this symmetrical three-tone, sine-wave testing of the opposite sideband. These results reinforce the analysis given above of the performance of a Decoder having these level difference characteristics.

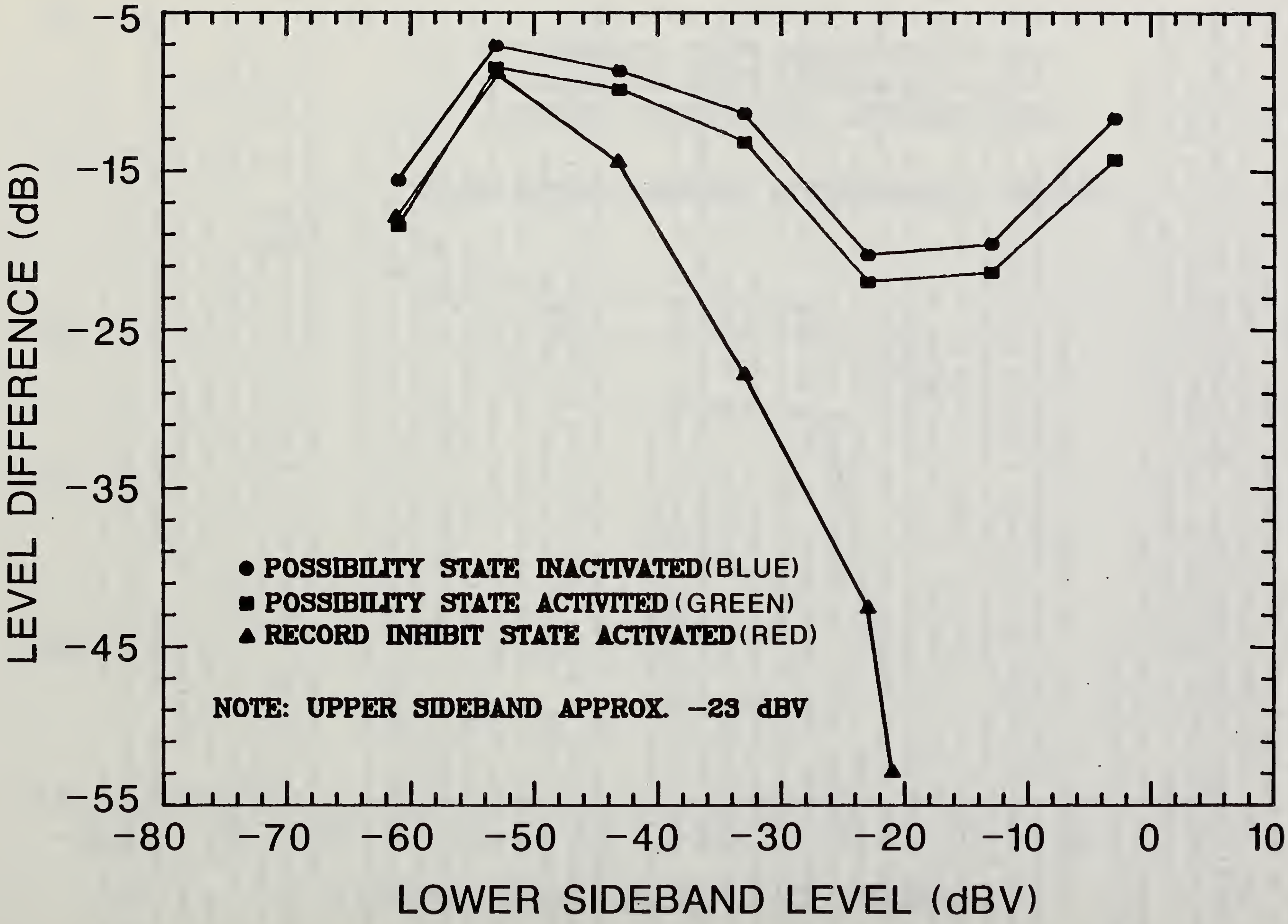


Figure 20 Difference level as a function of the lower sideband level (upper sideband level constant at about -23 dBV) that causes the POSSIBILITY AND RECORD INHIBIT states in the Decoder.

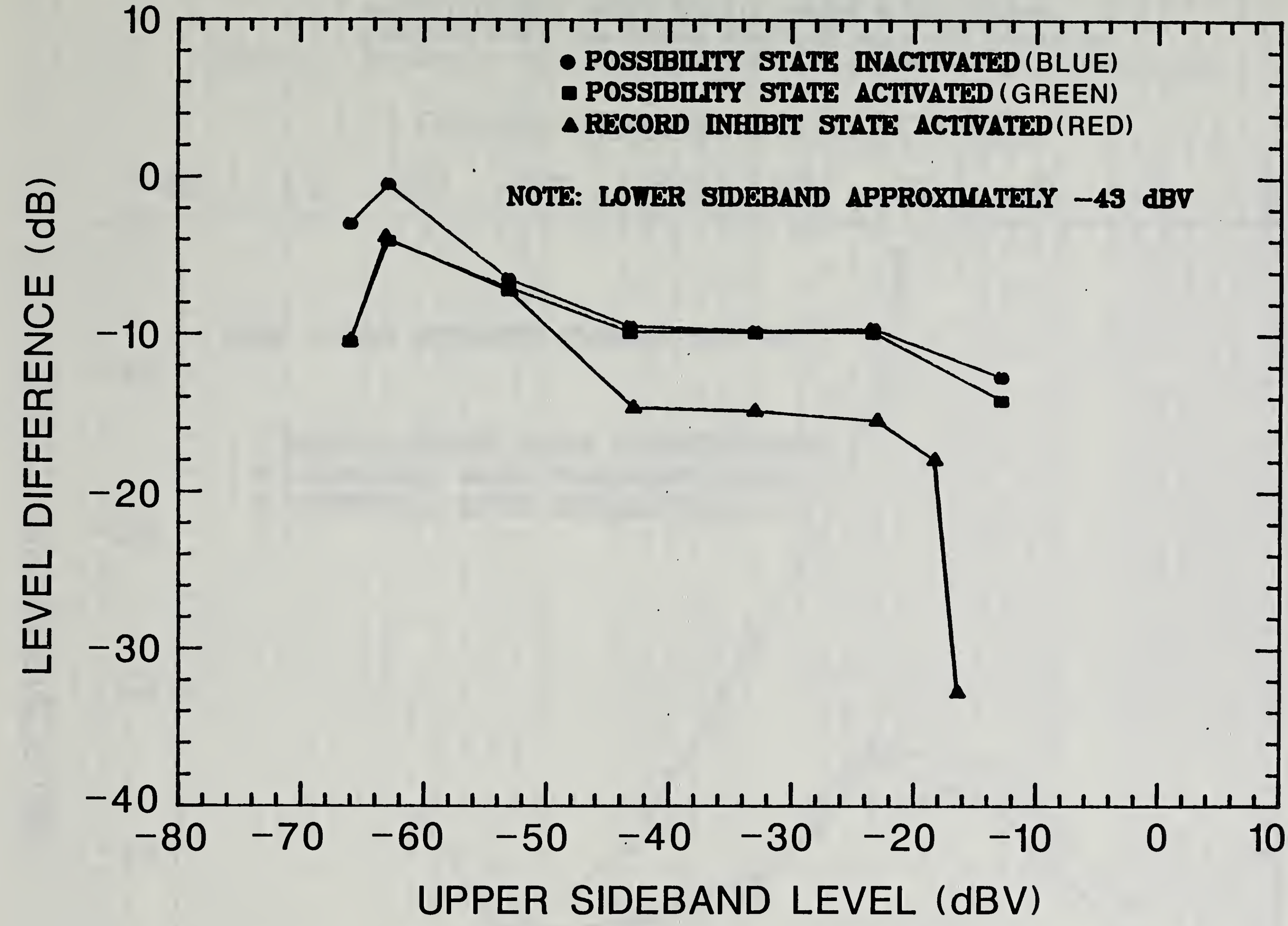


Figure 21 Difference level as a function of the upper sideband level (lower sideband level constant at about -43 dBV) that causes the POSSIBILITY AND RECORD INHIBIT states in the Decoder.

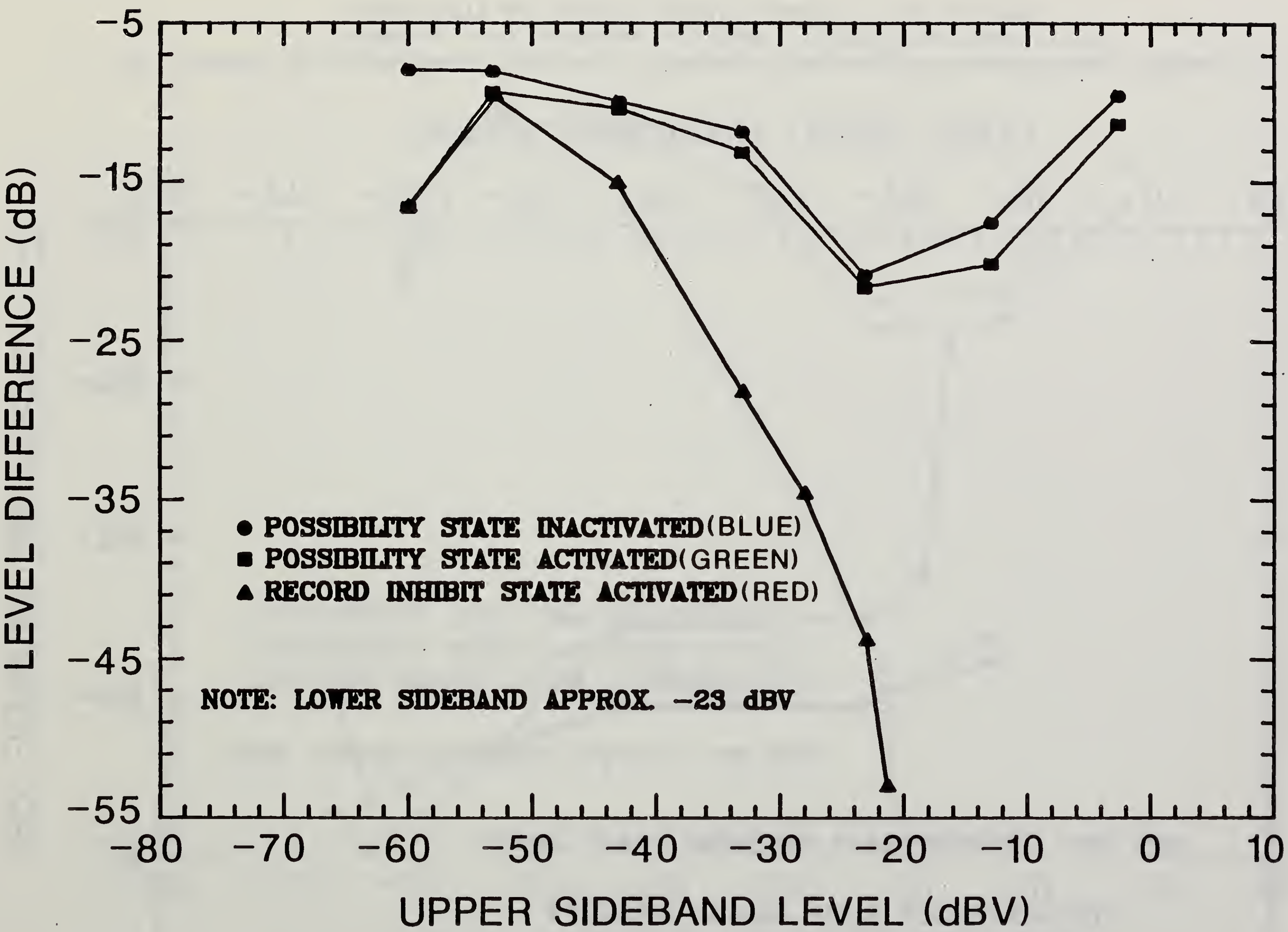


Figure 22 Difference level as a function of the upper sideband level (lower sideband level constant at about -23 dBV) that causes the POSSIBILITY AND RECORD INHIBIT states in the Decoder.

Functional Triangle, Ramp (Sawtooth), and Square-Wave Signals

Another test of the performance of the Decoder is to apply input signals with fundamental time-domain waveforms that may be representative of those found in many kinds of audio input signals. For this purpose, functional triangle, sawtooth (ramp), and square-wave input signals were utilized. These signals are fairly rich in harmonics of the fundamental frequency of the waveform; hence, the amplitude-frequency spectrum of these signals can be expected to potentially satisfy the conditions required for the Decoder circuitry to detect the presence of a notch.

In this case, a steady-state (fixed amplitude and fundamental frequency) functional waveform is applied while the scan controller of the Decoder is operated in the search mode (i.e., the center frequency of the frequency modulator is varied approximately 300 Hz about its nominal value). As with any kind of input signal, if the POSSIBILITY state exists during the sweep time of about 35 seconds, the scan controller stops the frequency modulator at that point, and the Decoder begins integrating the bandpassed signals. For these functional waveform tests, input signal levels of both relatively low amplitude (~-30 dBV) and high amplitude (~-10 dBV) were applied. Also, the fundamental frequency of these functional waveforms was varied from 200 Hz to 2000 Hz in (generally) 50 Hz intervals (see figure 23).

For triangle waveforms at the -30 dBV level, the amplitude of the spectral content in the notch region is relatively low, and no POSSIBILITY states occurred. However, at the -10 dBV level, the POSSIBILITY state and RECORD INHIBIT state were observed for triangle waves at 400 Hz and at 500 Hz. The POSSIBILITY state only was observed at 450 Hz. The RECORD INHIBIT state at 400 Hz and 500 Hz was obtained at input signal levels as low as ~-15 dBV and ~-20 dBV, respectively.

Square-wave input waveforms, which are rich in odd harmonics, caused POSSIBILITY/RECORD INHIBIT states to occur at both the -30 dBV and -10 dBV levels in the Decoder. For a -30 dBV square-wave signal level, RECORD INHIBIT was obtained for inputs with fundamental frequencies from 380 through 515 Hz. For a -10 dBV level, RECORD INHIBIT was observed with fundamental frequencies from 325 through 500 Hz and from 600 through 675 Hz. Figure 24 shows the frequency spectrum of a square wave input with a fundamental frequency of 486 Hz superimposed on the three bandpass filter characteristics plotted on a frequency scale from 0 through 5 kHz. Here, the 7th and 9th harmonics (identified by numbers next to the appropriate peaks) fit very near to the center frequency of the lower and upper sideband filters of the Decoder while the relatively small 8th harmonic (due to the non-ideal, square-wave signal) fits near the center of the notch passband. This spectrum is discriminated by the Decoder as having the notch characteristic required for causing a RECORD INHIBIT state.

For sawtooth (ramp) waveforms, which are rich in both odd and even harmonics, the occurrence of the POSSIBILITY and RECORD INHIBIT states was also prevalent at both the -30 dBV and -10 dBV levels. For -30 dBV inputs, RECORD INHIBIT was observed for fundamental frequencies from 700 to 900 Hz and 1050 to 1150 Hz. RECORD INHIBIT could also be obtained at 600 Hz for an input signal level of

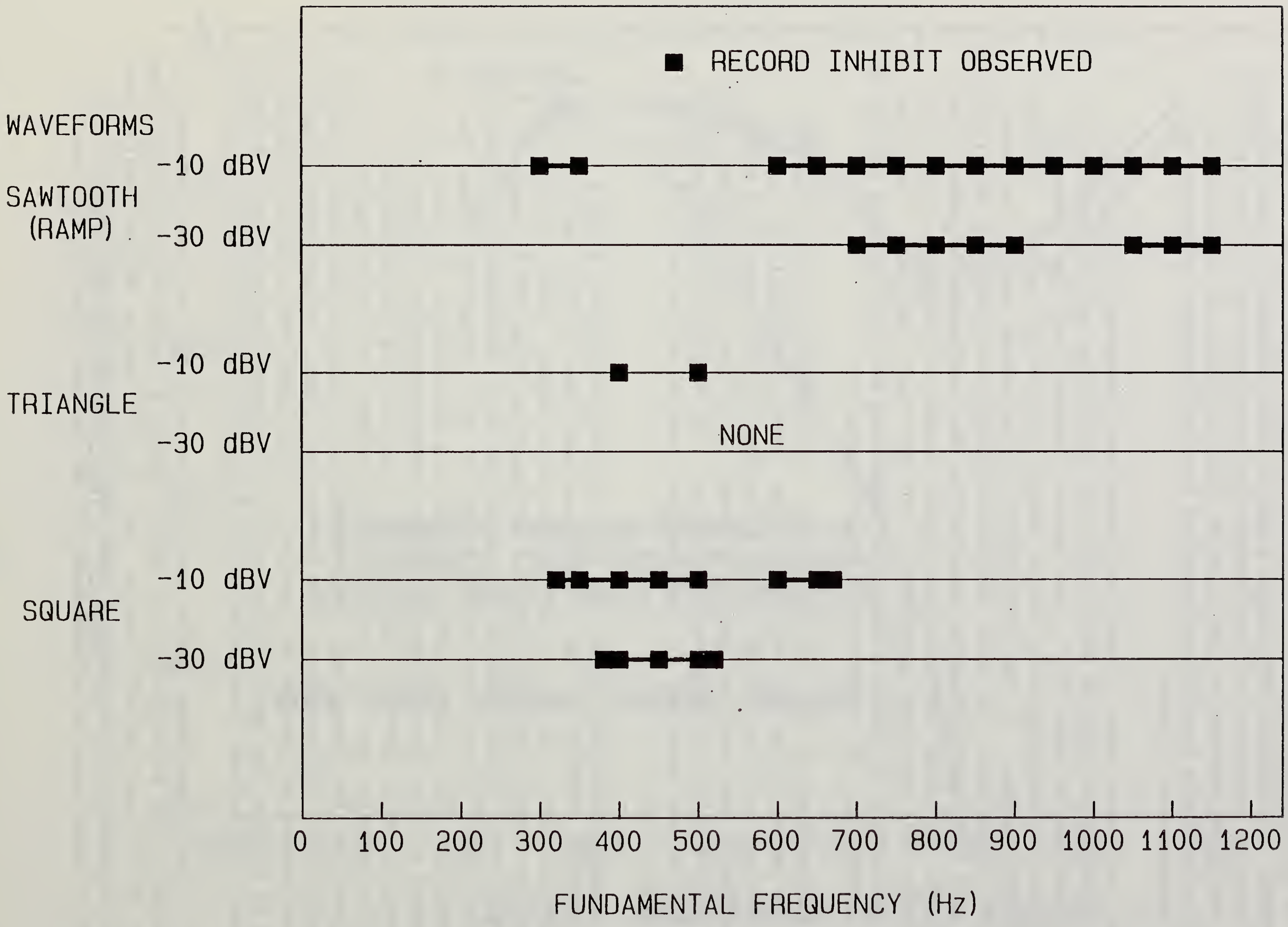


Figure 23 Summary of the RECORD INHIBIT activity observed for two different amplitudes of sawtooth (ramp), triangle, and square-wave input signals applied to the Decoder. Data points were taken at generally 50 Hz intervals from 200 Hz to 2 kHz. No RECORD INHIBIT states were observed above 1150 Hz.

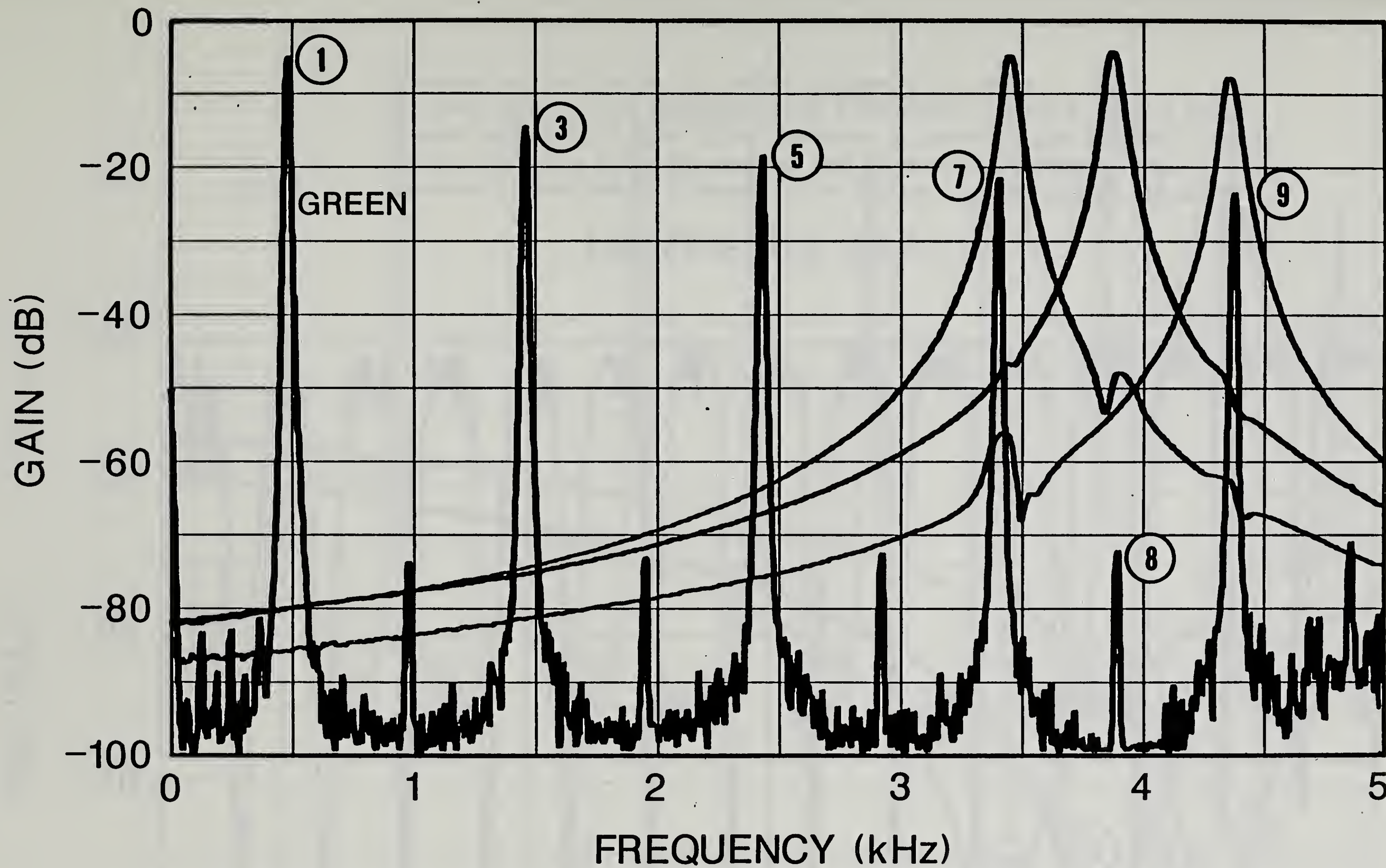


Figure 24 Frequency spectrum (in green) of a 486 Hz (fundamental) square wave input signal superimposed on the frequency response characteristics of the three bandpass filters in the Decoder.

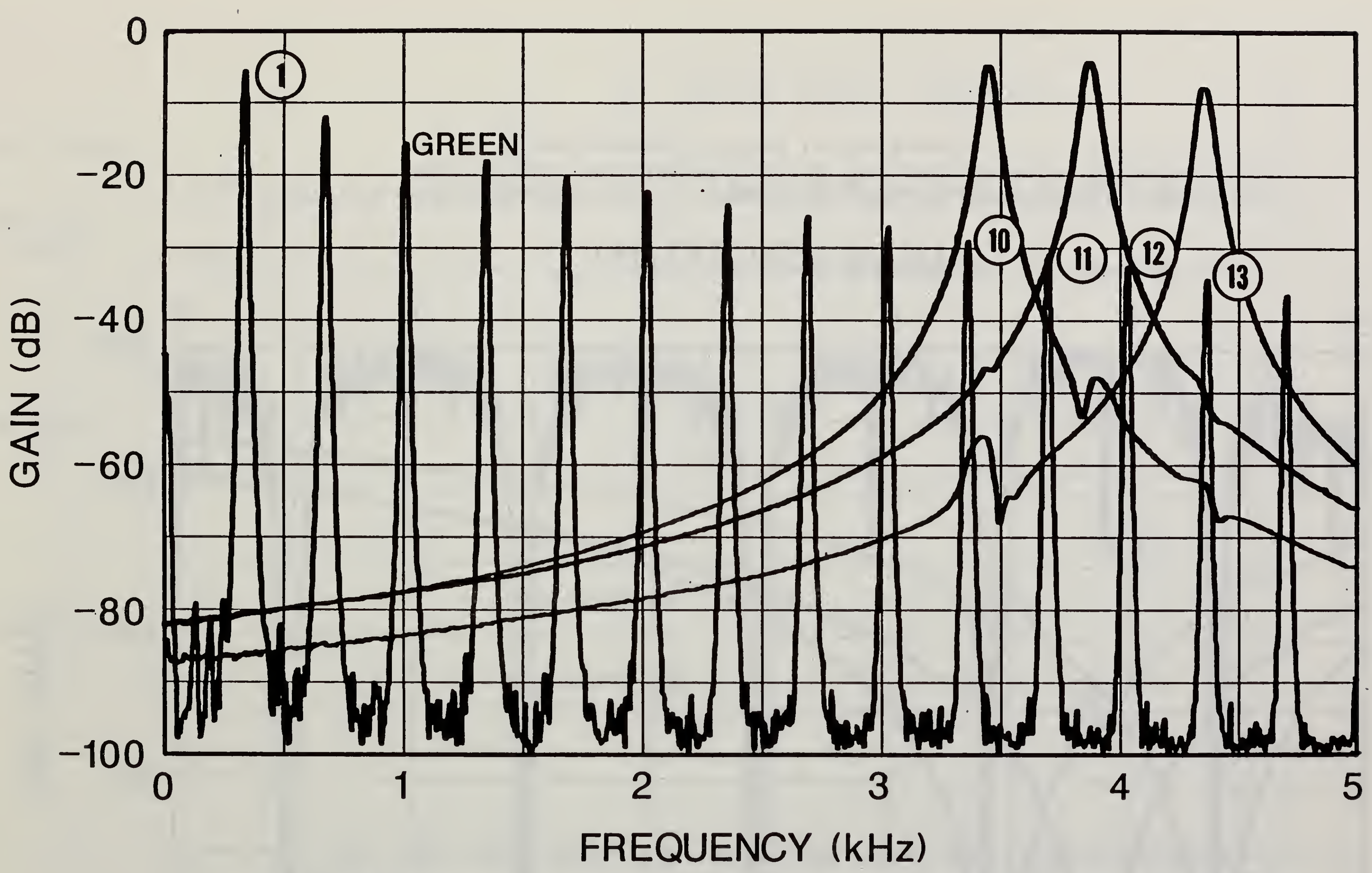


Figure 25 Frequency spectrum (in green) of a 337 Hz (fundamental) sawtooth (ramp) waveform superimposed on the frequency response characteristics of the three bandpass filters in the Decoder. Note the position of the 10th and 13th harmonics relative to the peak response of the lower and upper sideband filters.

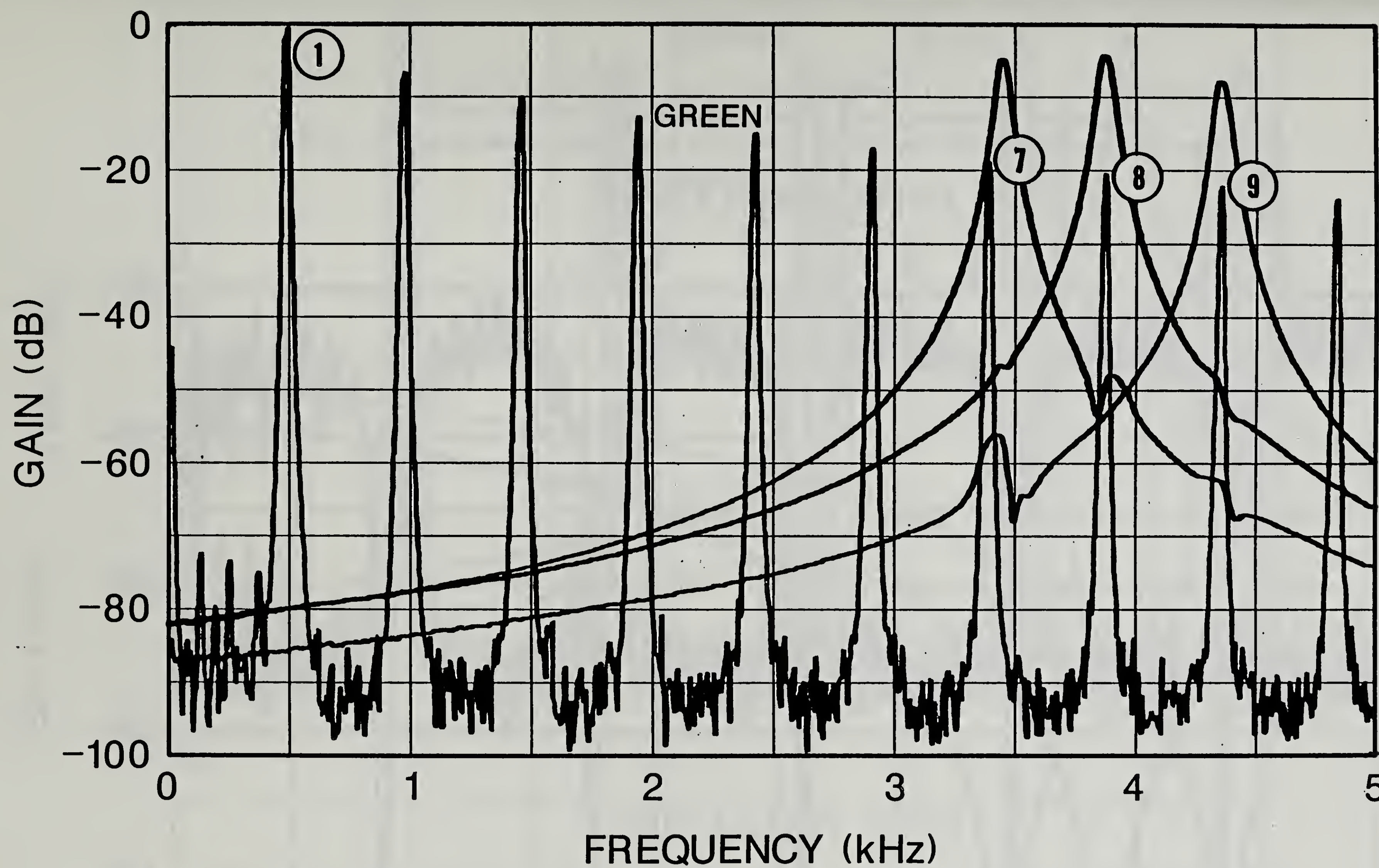


Figure 26 Frequency spectrum (in green) of a 486 Hz (fundamental) sawtooth (ramp) waveform superimposed on the frequency response characteristics of the three bandpass filter in the Decoder.

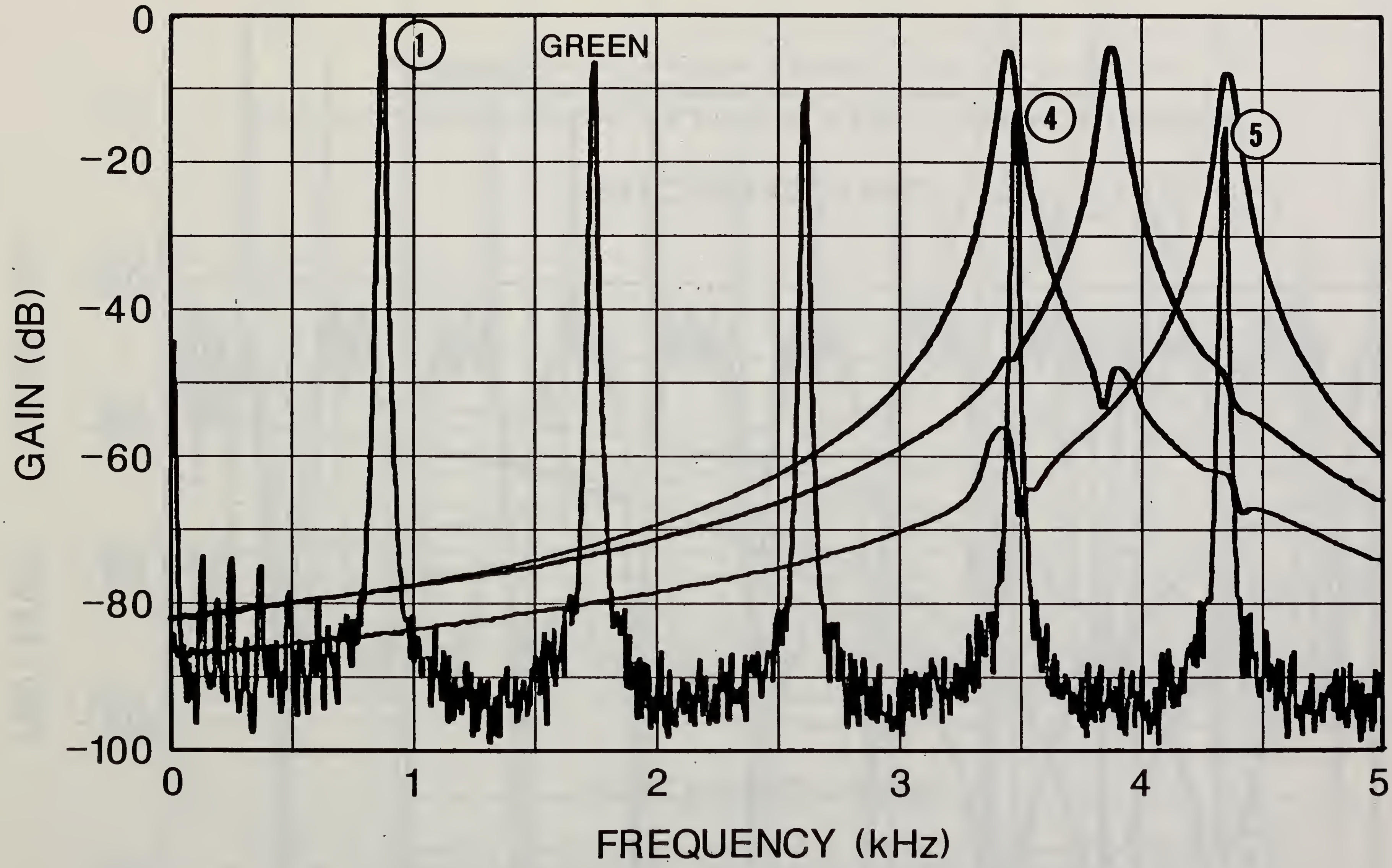


Figure 27 Frequency spectrum (in green) of an 870 Hz (fundamental) sawtooth (ramp) waveform superimposed on the frequency response characteristics of the three bandpass filters in the Decoder.

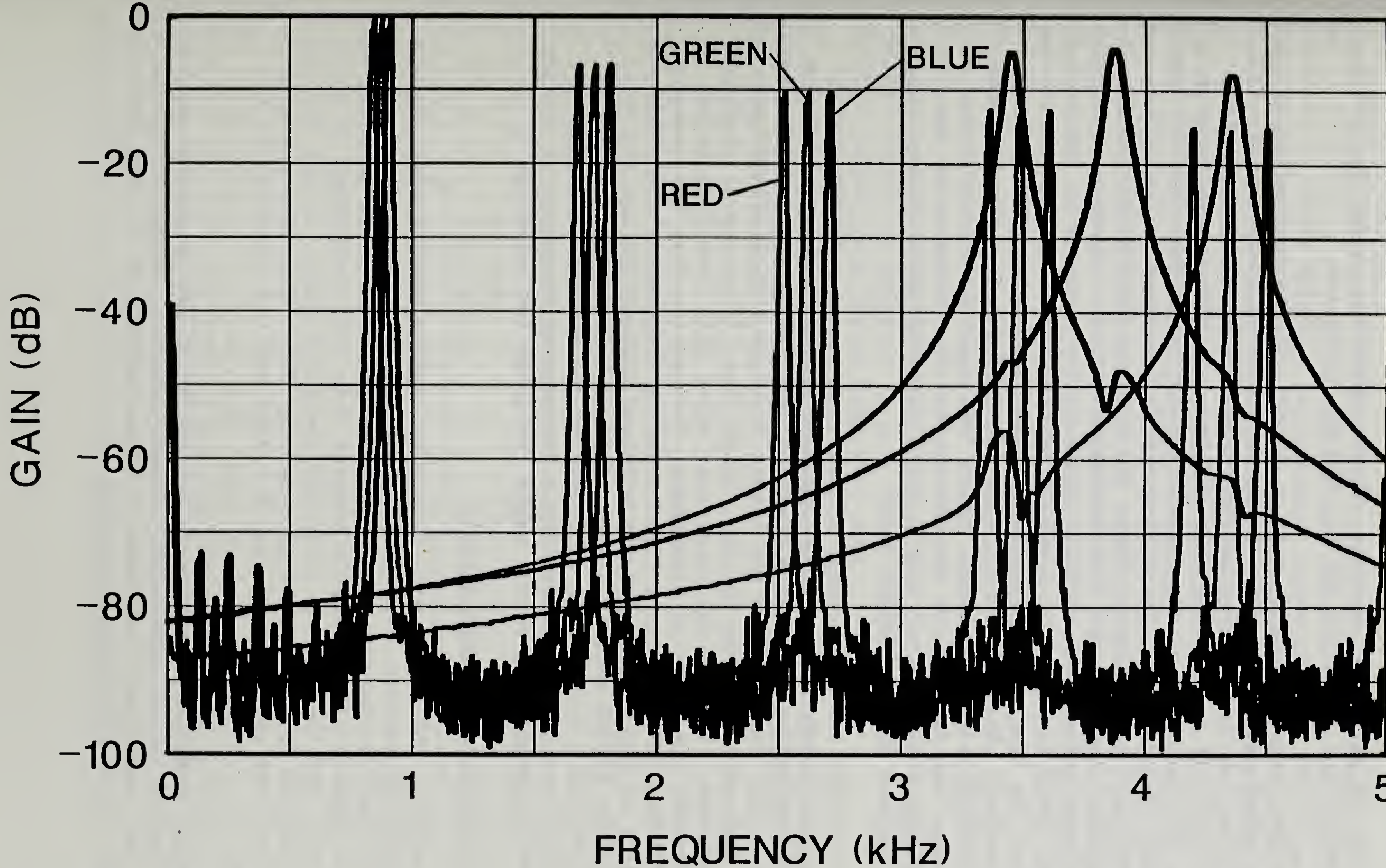


Figure 28 Frequency spectra of an 841 Hz (in red), an 870 Hz (in green), and a 901 Hz (in blue) sawtooth (ramp) waveform superimposed on the frequency response characteristics of the three bandpass filters in the Decoder.

~-19 dBV. At the -10 dBV level, however, RECORD INHIBIT occurred for 300 and 350 Hz input frequencies, and from 600 to 1150 Hz, inclusive.

Figure 25 shows the frequency spectrum of a -10 dBV, 337 Hz sawtooth waveform input signal superimposed on the three bandpass filter characteristics. The sawtooth signal is rich in both even and odd harmonics of its fundamental frequency at 337 Hz. It is apparent that the 10th and 13th harmonics of this sawtooth signal fit very near to the center frequencies of the lower and upper passband filters, respectively. The 11th and 12th harmonics are on either skirt of the notch passband filter. Thus, the 11th and 12th harmonics are highly rejected while the 10th and 13th harmonics are not. Consequently, this input signal is discriminated as having a notch in its spectrum at the proper frequencies, and the Decoder causes a RECORD INHIBIT state. Figure 26, on the other hand, shows the frequency spectrum of a sawtooth signal with a 486 Hz fundamental frequency. Here, the 7th , 8th, and 9th harmonics fit very near to the center frequencies of the lower, notch, and upper passband filters. All three harmonics are detected by the Decoder and discriminated as the spectrum of a signal that does not have a notch characteristic; hence, this signal did not cause a RECORD INHIBIT state.

The spectrum of another sawtooth (ramp) signal that caused a RECORD INHIBIT state is shown superimposed on the three passband filter responses in figure 27. The fundamental frequency of this sawtooth waveform is 870 Hz. The 4th and 5th harmonics fall very near to the center frequency of the lower and upper passband frequencies of the Decoder, while there is no energy component in the notch passband. To show the range over which this condition holds, figure 28 illustrates the range of fundamental frequencies that the sawtooth signal caused RECORD INHIBIT (with a fixed scanning frequency), i.e., how far down on the skirts of the upper and lower sideband filters the 4th and 5th harmonics caused the RECORD INHIBIT state. The red spectrum is for the sawtooth waveform with a fundamental frequency of 841 Hz while the blue spectrum is for the sawtooth waveform with a fundamental frequency of 901 Hz. The 870 Hz sawtooth waveform is shown, again, in green for reference.

Actual Encoded and Unencoded Audio Signals

During normal usage, a DAT system will likely be used to record audio signals from a wide variety of sources. For purposes of the NBS evaluation of the CBS-supplied DAT Recorder/Decoder, the variety of possible source material was narrowed down to commercially-available compact disc (CD) recordings of music. All of the CDs used for these tests contained unencoded music, except for one, which was an unencoded CD-player test disc made by CBS. This disc was used in helping to evaluate the operation of the CD player and the interface unit discussed below.

The table given in Appendix B contains a list of the 55 different CDs that were used in this testing of actual music source material; number 22 is the CBS test disc. All subsequent references to these discs will use the disc reference number for identification purposes.

As described below, the Decoder was tested using both unencoded and encoded music in order to determine whether any cases could be observed of false positive or false negative performance, respectively. The false positive and false negative nomenclature refers to the performance of the Decoder. During the playing of the music into the Decoder, the on-off activity (as a function of time) of the POSSIBILITY and RECORD INHIBIT lights (yellow and red LEDs) on the Decoder were recorded, as described below. If the Decoder indicated a RECORD INHIBIT during the playing of any part of a musical piece, this result was designated a "Positive" RECORD INHIBIT case. Similarly, if no indication of a RECORD INHIBIT was made during an entire piece, this result was designated a "Negative" RECORD INHIBIT case. Depending on whether the music was encoded or unencoded, the results of the tests either matched what was expected, and were designated a "True" case, or did not match, and were designated a "False" case. Thus, if unencoded music was played into the DAT Recorder/Decoder and the Decoder indicated no RECORD INHIBITs, this result is referred to as a "True Negative" case. Conversely, if a RECORD INHIBIT was indicated, this result is referred to as a "False Positive" case. For encoded music played into the DAT Recorder/Decoder, on the other hand, a RECORD INHIBIT result is referred to as a "True Positive" case, whereas a no RECORD INHIBIT result is a "False Negative" case.

To summarize, table 3 defines the four types of possible test cases with respect to the encoding of the source and the corresponding performance of the DAT Recorder/Decoder.

Table 3

The Four Types of Possible Test Cases

Decoder Indication	Source Material	
	Encoded	Unencoded
RECORD INHIBIT	True Positive	False Positive
No RECORD INHIBIT	False Negative	True Negative

The test setup for the unencoded music tests is shown in figure 29. Since both the CD player and the DAT Recorder/Decoder are designed to work with standard unbalanced IHF (Institute of High Fidelity) signal levels, the two (left/right) output channels of the CD player were directly connected to the input channels of the DAT Recorder/Decoder. The CD player used a cartridge which allowed up to ten CDs to be played automatically.

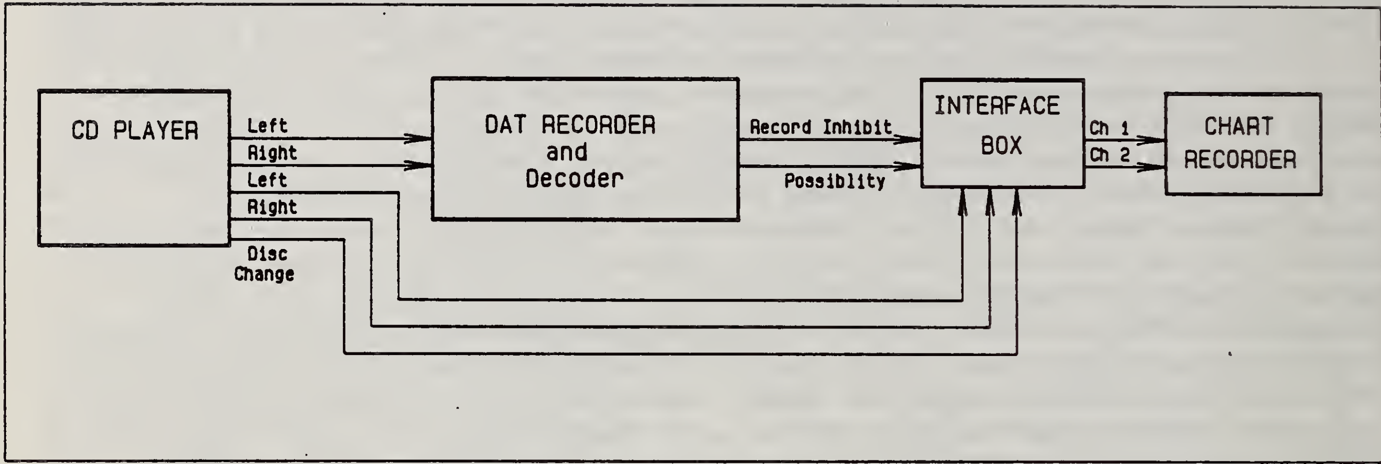


Figure 29 The equipment setup for running unencoded music tests

During the tests the operation of the CD player and the DAT Recorder/Decoder was monitored and recorded on the continuous roll of an oscillographic chart recorder. The Decoder logic signals from the yellow (POSSIBILITY) LED and the red (RECORD INHIBIT) LED were connected via optical isolators to an interface box and then to the chart recorder. The use of optical isolators minimized the effects of induced noise, ground loops, etc. caused by making external connections to the Decoder circuit. The two (left/right) audio channels and a third signal, which indicated when the CD player was changing discs, were also connected to the interface box. This box contained electronics to combine the POSSIBILITY and RECORD INHIBIT signals from the Decoder into one logic state signal, which was then monitored by the chart recorder. Similarly, the three signals from the CD player were combined and recorded on another channel. Thus, this dual-channel recording setup allowed the operation of the Decoder to be related to the specific disc, track, and time of occurrence within the track for any test run. Numerous overnight test runs of several hours length were thereby made unattended.

The setup for running tests with encoded music is shown in figure 30. Because the CBS-supplied Encoder was designed for use in a recording studio environment, it has relatively high level, balanced, input and output signals (different than the IHF standard). Thus, the Encoder should not be directly connected to the CD player and DAT Recorder/Decoder. Consequently, a commercially-available interface unit (the Match-Maker) was used to provide compatibility between the Encoder and the other equipment.

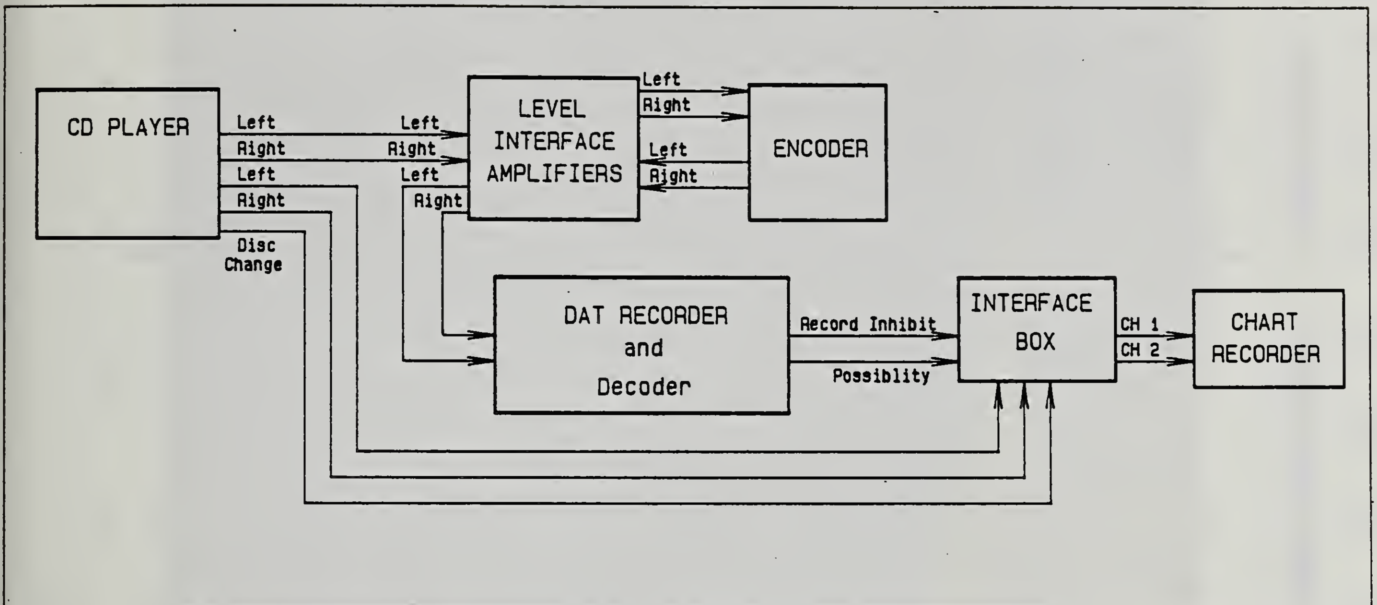


Figure 30 The equipment setup for running encoded music tests

Because the Decoder scans the three narrow passband filters (see section 2.3 on the operation of the DAT Recorder/Decoder), the results of the false positive and false negative tests tend to vary from run to run. The reason for this variability is that there is not a fixed relationship between the incoming music and the modulation frequency of the scanner. Also, any notches that are being detected are of finite duration. This finite duration has several causes. In the case of false positives, the notches that are being detected are those that can occur naturally in the frequency spectrum of the music. Because of the scanning in the Decoder, these naturally-occurring notches need not be at the nominal encoding notch frequency, but can be anywhere from about 3.5 to 4.1 kHz. The duration of a given notch in the spectrum is dependent upon the particular piece of music.

Two examples of music with naturally-occurring notches are shown in figure 31. This figure shows spectrographs of two selections (disc numbers 17 and 34 from the table given in Appendix B) which caused false positives. The spectrograph is a three axis plot with time shown on the horizontal axis and frequency shown on the verticle axis. For these plots, the time is about one minute and frequency is from 0 to 6 kHz. The level within each frequency band is indicated by color; a reference color bar shows the relative levels from 0 to -80 dB. Spectrograph 31 A shows one minute of music from the Star Wars Proto, starting at about one minute into the piece. Spectrograph 31 B shows the first one minute of music from Mendelssohn's Wedding March. False positives were observed for the Star Wars Proto at about 1:35 into the piece and at about 0:35 into the Wedding March. Thus, for both spectrographs the false positives occurred near the center of the time scale.

In the case of false negatives, the notches to be detected are purposely provided by an encoder. These notches are also of varying duration. As mentioned in section 2.1, the CBS-supplied Encoder produces notches in only part of the encoded music. Also, as described in section 2.3, the comparison

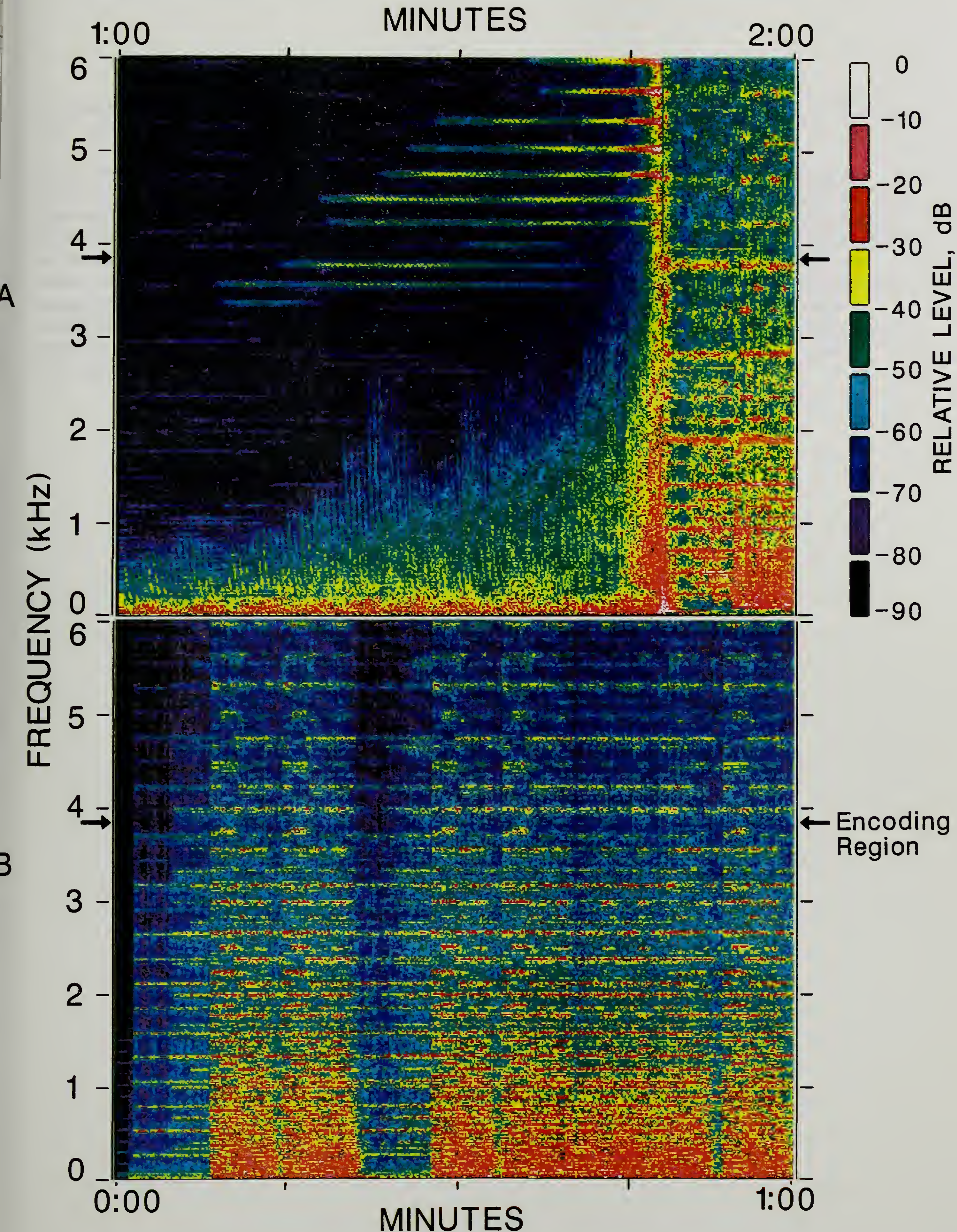
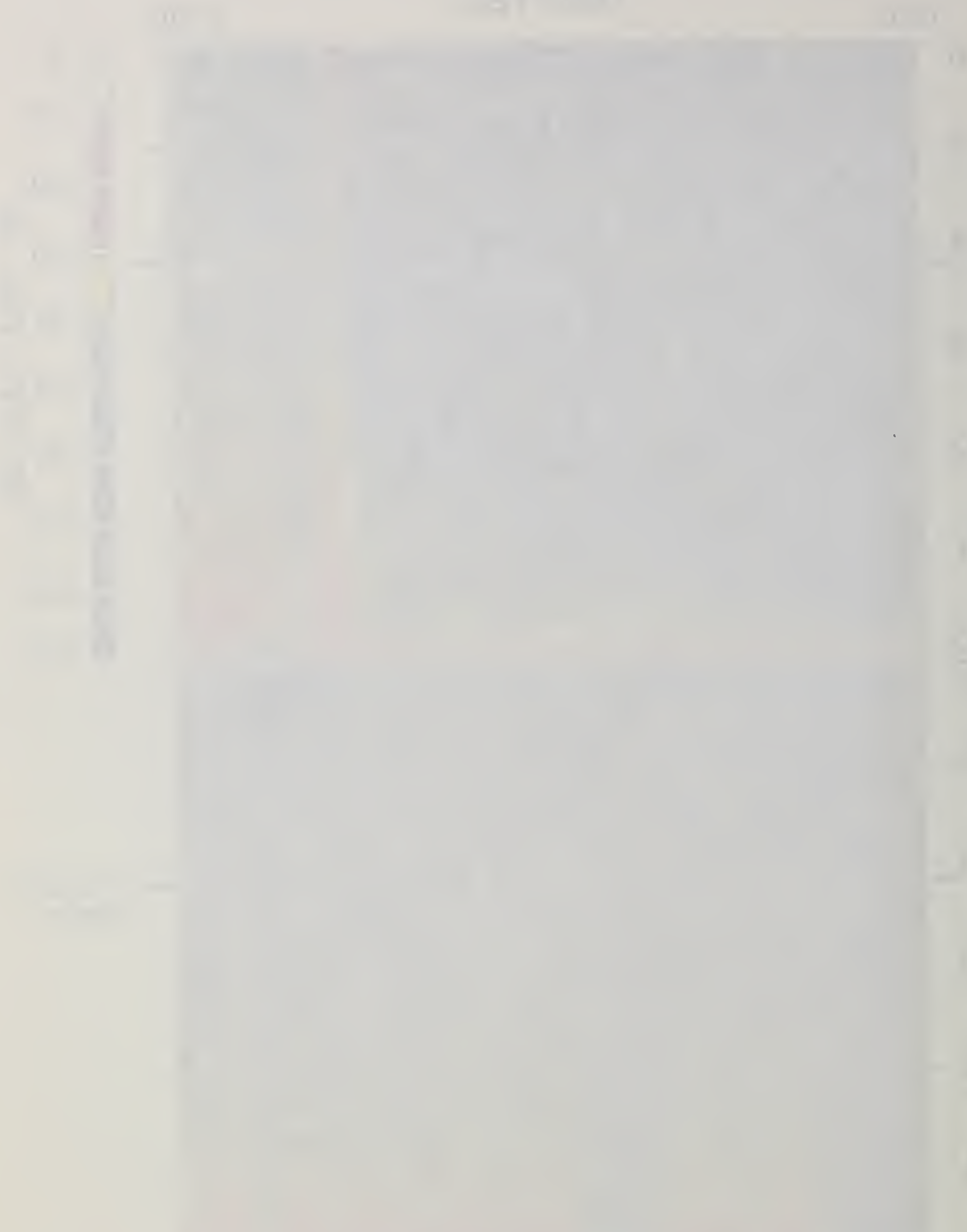


Figure 31 Color spectrographs of music with naturally-occurring notches which caused False Positives. Spectrograph A shows one minute of music from the Star Wars Proto starting at 1:00. Spectrograph B shows the first one minute of Mendelssohn's Wedding March.



process in the Decoder is inactivated if the input music levels are too low at the sideband frequencies.

With the ~35 second scanning period of the Decoder, it can take the Decoder from 0 to about 17 seconds to begin identifying the presence of a notch in the incoming signal spectrum. Then, the notch must be present for an additional 13 to 15 seconds in order to cause a RECORD INHIBIT. Hence, notches that have a duration longer than about 15 seconds will be detected by the Decoder on some runs but not on others. The longer the duration of the notch, the more consistently the Decoder was observed to recognize it and cause a RECORD INHIBIT. The cases of false positive and false negative performance that were measured during the NBS tests are summarized in sections 3.2 and 3.3 that follow.

3. DOES THE SYSTEM ACHIEVE ITS PURPOSE?

3.1 Interpretation of the Question

This question has been interpreted to mean that the copy protection code system will achieve its purpose if it inhibits the CBS-supplied DAT Recorder/Decoder from recording all or parts of an encoded signal and does not inhibit the recording of unencoded signals. A signal is considered encoded if it has passed through the Encoder; otherwise, it is unencoded.

3.2 Decoder False Positives

As described in section 2.4 above, the tests for cases of false positive performance occurring in the Decoder were carried out by recording the POSSIBLE and RECORD INHIBIT states of the Decoder when the DAT Recorder/Decoder was receiving unencoded music signals. This section summarizes the results of the test setup and measurements described in section 2.4 under Actual Encoded and Unencoded Audio Signals.

Because of the scanning operation within the Decoder, the RECORD INHIBIT indication of the Decoder varies from one playing of a compact disc (CD) to another playing of the same CD (see discussion in section 2.4 above). Therefore, this behavior required playing the music pieces several times to determine the random possible Decoder responses. The false positive testing was done in two phases, as described below. All 54 of the music discs listed in Appendix B were used to test the Decoder for possible false positive operation. The Decoder responses obtained are reported both in terms of an entire CD and in terms of the individual tracks on the discs. The 54 CDs tested had a total of 502 tracks.

During the first phase of the false positive tests, all of the tracks on the discs were played into the Decoder. Most of the CDs were played two or three times; five of the 54 were played only once, and two were played more than three times. The Decoder indicated a RECORD INHIBIT on ten of the 502 tracks. These ten tracks were on five different CDs. Table 4 on the next page shows which discs indicated false positives and the number of passes for each disc. Thus, with only a few passes, about two percent of the tracks or nine percent of the CDs had a false positive indication on at least one of the playings.

Table 4
False Positive Data

Disc Number	Tracks	Full Disc Run: Number Of Runs	Full Disc Run: False Positive Tracks	Partial Disc Runs: Track Number (Times)	Partial Disc Runs: Number of Runs	Partial Disc Runs: False Positive Rate
3	6	6	1	4	21	10%
11	10	3	0	6 (3:30-5:30)	24	38%
				6	48	25%
14	10	3	0	10 (0:00-1:30)	10	20%
				10	48	8%
17	10	2	0	1 (0:00-2:30)	79	34%
				1	47	11%
25	12	2	1	5 (0:00-2:30)	8	12%
				5	47	62%
				8 (2:30-5:00)	52	27%
				8	47	43%
26	4	2	1	1	47	11%
33	16	3	4	4	21	33%
				10	21	14%
				11	21	0%
				15	21	48%
34	9	3	3	1	21	29%
				5	21	29%
				9	21	86%
48	3	2	0	1 (15:30-18:00)	9	11%
52	11	2	0	5 (0:00-1:30)	14	29%
				6 (0:00-2:00)	13	8%

The second phase of the false positive tests examined individual tracks or parts of tracks, using a higher number of passes than were run for the complete CD tests of phase one. The tracks or parts of tracks were selected based on the POSSIBILITY indication of the Decoder observed during phase one. To become a RECORD INHIBIT, the POSSIBILITY indication must last for about 15 seconds. Thus, tracks or parts of tracks which had relatively long POSSIBILITY states of about 8 to 14 seconds, were tested further. The selected tracks and track times are also listed in table 4, along with the number of runs and the percent rate of false positive occurrences. With this additional testing, false positive cases were found on a total of 16 different tracks on ten of the CDs. Thus, over 3 percent of the total tracks tested and 18 percent of the CDs tested had cases of occasional false positives.

3.3 Decoder False Negatives

The tests for cases of false negative performance were carried out by recording the state of the POSSIBLE and RECORD INHIBIT indications of the Decoder while it was receiving encoded music signals. This section summarizes the results of the test setup and measurements described above in section 2.4 under Actual Encoded and Unencoded Audio Signals.

The test signals were encoded by one of the two CBS-supplied Encoders, designated Encoder 006 or Encoder 012 (corresponding to their serial numbers). Eighteen CDs with 150 tracks were tested for false negative cases using Encoder 006, and 32 CDs with 280 tracks were tested for false negative cases using Encoder 012.

The false negative test data is summarized in table 5, which identifies the discs by their reference number, and gives their associated number of tracks. In general, each disc was run with the test setup either two or three times. Seven of the discs (numbers 39 - 45) were tested for higher numbers of runs. For each track on a disc, the track is labeled as a False Negative Track if no RECORD INHIBIT was observed during at least one of the runs. Table 5 gives the number of false negative tracks for each disc that was tested using the two different encoders. If (during at least one of the runs) no RECORD INHIBITs were observed for the entire disc, the disc was labeled as showing a disc false negative (DFN), and the disc is so marked in table 5.

Table 5
False Negative Data

Disc Number	Tracks	Encoder 006		Encoder 012	
		Number of Runs	False Negative Tracks	Number of Runs	False Negative Tracks
1	9	3	6	2	1
2	5	2	4	2	2
3	6	5	DFN	4	4
4	5	2	DFN	2	3
5	7	2	6	2	2
6	12	2	DFN	2	9
7	5	2	4	2	2
8	9	-	-	3	0
9	13	2	12	2	1
10	12	2	12	2	6
12	11	2	DFN	2	0
13	9	3	DFN	2	0
15	11	-	-	3	4
16	8	-	-	3	0
18	8	-	-	3	4
19	10	-	-	3	1
20	4	-	-	3	DFN
32	8	3	DFN	2	DFN
33	16	3	14	2	11
34	9	3	8	2	8
35	7	3	DFN	2	1
36	3	3	DFN	2	DFN
37	4	3	4	2	1
38	9	3	8	2	6
39	11	-	-	7	8
40	8	-	-	7	0
41	9	-	-	7	6
42	8	-	-	7	0
43	10	-	-	6	0
44	10	-	-	4	3
45	12	-	-	4	9
46	12	-	-	3	5

The results given in table 5 show a significant difference between the two Encoders in the false negative data obtained. As mentioned in the description of the encoders in section 2.1, Encoder 006 encoded less of the source material than did Encoder 012. A consequence of this difference is that the overall copy protection system (Encoder and DAT Recorder/Decoder) showed a much higher false negative rate when using Encoder 006 than it did when using Encoder 012. For example, in using Encoder 006, eight DFNs out of 18 discs tested were obtained, while in using Encoder 012 only three DFNs out of 32 discs tested were obtained. In terms of tracks, using Encoder 006, the system had 139 false negative tracks out of 150 tracks, and, using Encoder 012, the system had 112 false negative tracks out of 280 tracks tested. Thus, in only a few passes, using Encoder 006, the system did not inhibit the recording of ~93 percent of the tracks at least once, and, using Encoder 012, the system did not inhibit the recording of ~40 percent of the tracks at least once.

4. DOES THE SYSTEM DIMINISH THE QUALITY OF THE PRERECORDED MATERIAL?

4.1 Interpretation of the Question

This question was interpreted as asking whether the quality of the recorded material is diminished as revealed by objective electrical measurements, subjective listening tests, or both. These measurements and tests would apply to the overall performance of the recording and playback system. Electrical measurements include those parameters that are commonly used to specify the overall performance of high-quality audio systems. Subjective tests address whether or not discriminating listeners can detect the presence or absence of encoding.

4.2 Effect on the Performance of an Audio System

The electrical components of a modern high-quality audio system are typically designed and built to meet rather tight performance specifications with regard to their frequency response. For example, a compact disc player or a good power amplifier typically would have an amplitude-frequency characteristic that is flat to within within ±0.5 dB over the frequency range from 20 Hz to 20 kHz. When it is specified, the phase response of a good power amplifier would typically be flat to within ±10 deg. Studio quality recording equipment would be expected to approach or exceed these same frequency characteristic specifications.

As indicated in sections 2.1 and 2.2, the Encoder has a notch in its amplitude-frequency characteristic that is more than 80 dB deep near 3840 Hz and has a very irregular phase response in that frequency region as well. Thus, if the Encoder were to be introduced into the music recording and reproduction process, there would be no question but that the frequency characteristics of the signal coming out of an amplifier in one's home would be degraded relative to the signal that was originally recorded prior to copy-prevention coding, in that the usual statements of system performance as being within ±X dB (amplitude) and ±Y deg (phase) would no longer be met.

It is essential, however, not merely to look at physical performance specifications but to address the question of the audibility of the encoding process. That question is examined in the remainder of section 4.

4.3 Subjective Response Testing Procedure

Professor Irwin Pollack of the University of Michigan served as a consultant to NBS on the design and implementation of the subjective listening tests that were carried out. His report to NBS, which is reproduced as Appendix C, discusses the philosophy and rationale behind the design of the tests and summarizes some of the test findings.

Two separate but related series of listening tests were carried out in which listening subjects attempted to discriminate between music which was presented to them "directly" (not encoded) and the same musical selection presented to them "encoded" (having passed through one of the two Encoders described in section 2.1). These two series of tests are briefly described below.

The musical selections presented to the subjects were selected from material suggested by the project sponsors (see Appendix D) or were selected by NBS staff. The selection and characterization of the music used in these studies are discussed in section 4.4 and in Appendix E.

The encoded and unencoded music selections presented to the subjects were prepared so as to be identical except for the effect of the Encoder. For both types of material, digital signals were converted to analog, passed through either the Encoder or a straight wire, and then converted back to digital format. For a given musical selection, every appearance of that selection with a given encoding state (direct or encoded) on the tape was a digital copy of the same material so that, for example, the encoded version of a given selection was always identically the same. The details as to how this tape preparation was accomplished are given in Appendix F.

The material was presented to the listening subjects using a professional quality digital audio tape recorder, audio processor, and high quality earphones or loudspeakers. The audio systems that were used are documented in Appendix G.

Serial Listening Test

In the listening tests that are hereafter referred to as "serial listening tests," short segments of music were presented sequentially to multiple listeners in a fixed sequence that was established by NBS staff. The music was presented as pairs that were repeated three times in close succession, with each member of the pair being either direct (D) or encoded material (E); for example, the presentation might be as follows:

Selection 1	D	E	D	E	D	E	(Different)
Selection 2	E	E	E	E	E	E	(Same)
Selection 3	E	D	E	D	E	D	(Different)
.	.	.	.	.	.	.	
.	.	.	.	.	.	.	
.	.	.	.	.	.	.	
Selection 2	D	D	D	D	D	D	(Same)
.	.	.	.	.	.	.	
.	.	.	.	.	.	.	
Selection 1	E	D	E	D	E	D	(Different)

The subjects were asked to decide whether the two members of the pair were the same or were different. Over the course of the test period (approximately 2 hours) 24 musical selections were presented in all four of the possible pairings (D-D, D-E, E-D, and E-E) with, as stated above, each presentation consisting of three repeats of a given pairing. Thus, the subjects were asked to make 96 choices as to whether the two members of a given pairing were the same or different. For each of the four sets of 24 selections, the order in which the selections appeared was randomized. For a given selection, the order in which the four possible pairings appeared was also randomized.

In addition to the musical selections, the presentation tape included a narrative description of the nature of the test and instructions to the

subjects. The text of this narrative is included in Appendix H.

The material was presented simultaneously to up to seven subjects, five using earphones and two using loudspeakers in a specially prepared listening room. The majority of the listening subjects consisted of volunteers who responded to a request that was sent out to the local chapter of the Audio Engineering Society or to an announcement that was sent to radio stations and recording studios in the Washington-Baltimore metropolitan area. Most of these subjects were individuals who are actively involved in music production, recording, or broadcasting. Additional subjects were selected from NBS employees, not involved with this study, or from friends who were known to be discriminating listeners.

Parallel Listening Test

In the listening tests that are hereafter referred to as "parallel tests," ten segments of music, each of about 1 minute in length, were presented to one subject at a time, over earphones. The selections were recorded synchronously on three parallel pairs (stereo) of tracks on a 24-track digital tape recorder. One pair of tracks (REF) was a reference pair, where the music was always presented directly, without encoding. On a second pair of tracks (A), the material was either encoded or direct, while on a third pair of tracks (B), the material was either direct or encoded, with one of the two pairs (A or B) always being encoded material and the other being direct material. The track pair on which encoded material was recorded was randomly assigned. Each subject could rewind the tape recorder and switch among the three pairs of channels at will. His task was to attempt to identify, for each musical selection, the track pair containing the encoded material.

The instructions to the subjects are included in Appendix H.

The subjects used for the parallel presentation study were sound recording engineers, audio equipment designers, or musicians.

4.4 Selection and Characterization of Recorded Material

The material used for the listening tests was taken from compact discs and from a Kurzweil Model 250 digital sampling keyboard (synthesizer) played by a professional musician in the NBS laboratory. Appendix E provides a list of the material selected and gives a description of some of the effects that were noted as likely to be caused by the encoding process. It is important to note that the effects described below and in Appendix E were observed with the aid of extensive physical measurements, which assisted the NBS staff member who selected the material to focus his listening attention on very specific musical passages or notes.

As indicated above, the compact discs used were chosen from lists (reproduced in Appendix D) supplied by the Home Recording Rights Coalition (HRRC) and by the Recording Industry Association of America. Additional material was selected based on physical measurements and listening by NBS staff, as described below. The list furnished by the HRRC supplied tracks and times at which audible effects had been noted by their consultant, using signal

processing based on his understanding of the encoding process. In the majority of these cases, the logic used by the Encoders supplied to NBS prevented the production of the encoding notch at these times. However, the encoding notch was produced for many examples cited by the HRRC.

Twenty musical items from compact discs were chosen for the serial presentation tests, along with four produced by the synthesizer. The selection was based upon real-time spectrographs of the music, and upon careful listening by the NBS staff. The material was selected so that enough musical components fell in the encoding region to result in a potential for detection. (The real-time spectrographs used to aid in material selection were obtained using an Hewlett-Packard (HP) 3565S Analyzer, and an HP 330 Computer. Spectrographs used to document the actual material used in the listening tests were obtained with a different HP analyzer and are discussed in Appendix E).

Ten of these compact disc selections were chosen for the parallel presentation test, with the length of each selection being increased from a few seconds to between 29 and 75 seconds.

In order to classify the possible results of encoding music, six possible effects were hypothesized, with considerable overlapping among effects. These hypotheses, based on the spectrographs and on extensive listening by NBS staff, are summarized below.

1. The deletion, or significant reduction in loudness, of a particular note. This might happen when the fundamental frequency or a predominant harmonic of a note occurs within the band of frequencies removed by the Encoder.

2. A decrease in brightness in the timbre of notes or chords. This effect would occur when considerable harmonic content falls in the encoding region, but with enough energy at other frequencies that the apparent loudness of the sound is not changed significantly. Effects 1 and 2 could shade into each other.

3. A change in the phenomenon sometimes called "chorus" of the string section of an orchestra. The instruments are not playing precisely in tune, and individual players are using varying amounts and rates of vibrato. As a result, the spectrum of a note or chord will show a spread over a considerable range of frequencies for a particular harmonic, instead of the energy being concentrated at one frequency. The encoding method under study can remove a portion of this spread. The Encoder logic is not very effective at preventing this sort of change in the signal, which would be expected to be more subtle than those noted in 1 and 2. Effects 2 and 3 could overlap each other.

4. A change in the character of transient sounds. This effect would be expected to be the most subtle and subjective of all effects noted. Such a change would affect the "impact" produced by percussive sounds, or result in a slight coarsening of the overall sound.

5. Changes in the nature of a sound due to switching of the Encoder from one

state to another (and perhaps back again) within one note or chord due to a slight change in spectral content.

6. Amplitude modulation of a harmonic of a trill of a single instrument or voice, which extends into the encoding region or through it.

The selections of music, in general, were made in order to explore the audibility of rather specific effects upon specific passages. Even more subtle overall quality changes may occur, but if they do, extended listening in a familiar environment may be required for evaluation.

The recorded material was characterized by obtaining spectrographs of each musical sample. This was done for both the serial presentation material (24 short samples) and the parallel presentation material (10 extended samples). As discussed at the end of section 2.4, the spectrographs show time on the horizontal axis and frequency from 0 to 6 kHz on the vertical axis. Relative signal level within each frequency band is indicated by color with the level decreasing as the colors progress through the spectrum: white, pink, red, yellow, green, cyan, blue, purple, and black.

Figure 32 shows the color spectrographs of the direct and encoded material for four of the short selections used in the serial presentations. For each pair, the first (left) spectrograph represents the direct material while the right-hand spectrograph corresponds to encoded music. The arrows just below 4 kHz indicate the encoding region. In figure 32, the Debussy selection shows that a strong harmonic of the violins is removed. The encoder ceased notching in the middle of the Barber selection.

Figure 33 shows the spectrographs of the direct and encoded material for the Bizet selection used in the parallel presentations, with the lower spectrograph corresponding to encoded music. This material is seen to be encoded most of the time, resulting in partial removal of a harmonic of a solo violin.

4.5 Results of Listening Tests

The results of the subjective listening tests show that, while the effects of the encoder are fairly subtle for some musical selections, there are some selections for which the subjects detected differences between direct and encoded material. The data indicate that the ability to hear effects of the encoder varies substantially between individual listeners and, especially, between musical selections.

Serial Listening Test

A total of 87 subjects took the serial listening test -- 61 used earphones, 23 listened over loudspeakers, and 3 more used a "mixed" listening mode in which they took part of the test using earphones and listened over the loudspeakers for the remainder. Generally speaking, the earphone group achieved higher

discrimination scores than the loudspeaker group. The tests were not designed to evaluate differences between earphone and loudspeaker listeners in sensitivity to Encoder effects. Whatever differences may exist do not preclude using both groups to test the general hypothesis of no audible difference between direct and encoded music. In fact, both groups provide evidence that the difference is audible.

The design of the serial listening test ensured that a subject who persistently answered either "different" or "same" to most of the test items could not, by so doing, achieve a high score. This is because of the balance in the way the musical selections were presented: each selection was presented in all four possible combinations: DD, DE, ED and EE (where D = direct, or unencoded, and E = encoded). As an extreme example, a subject who answered "different" for every test item would get exactly half of the items correct (the DE and ED items) and the other half wrong (DD and EE), for a total score of exactly 50% correct.

In summarizing the test results, it is of interest to consider the average scores (percent correct discrimination) obtained for individual musical selections and for listeners.

Average Scores for Musical Selections

Table 6 shows the average correct discrimination scores, in percent, for each of the 24 musical selections in the study, averaged across all 87 listening subjects. The selection names are defined fully in Appendix E. Since each selection was presented four times (DD, DE, ED and EE) in the serial test, each average is based on $87 \times 4 = 348$ responses to the test items. Among the 348 responses there were 2 instances of missing data, where the listener did not mark either "D" or "S" in response to a test item. These cases were scored as incorrect.

Of the 24 selections in table 6, eight have average scores that exceed 50% by a statistically significant margin ($p < 0.025$). Statistical significance is evaluated by calculating the probability of obtaining a score as high or higher than the observed score on the assumption that the listeners could not hear any difference but were only guessing -- with a 50% probability of guessing the correct answer -- on each of the 348 responses. For example, using the binomial probability distribution, one can calculate that the probability is 0.95 that the observed percent of correct responses will be within ± 5.4% of 50% if 348 responses are obtained by guessing. In all, 9 of the 24 musical selections differ from 50% by more than 5.4%. This is substantially more deviation from 50% than could be explained by chance, since only about 1 or 2 selections (about 5%) should fall outside the interval 50% ± 5.4% if the responses were all guesses. Of the 9 selections with scores outside the central 0.95 probability range, 8 were above that range (i.e. above 55.4%) and the remaining 1 was below that range (less than 44.6%).

The most pronounced deviation above 50% was obtained for the S1 selection, which was correctly discriminated 90.5% of the time in the test. This selection consists of an F_7-B_7 chord created on a music synthesizer

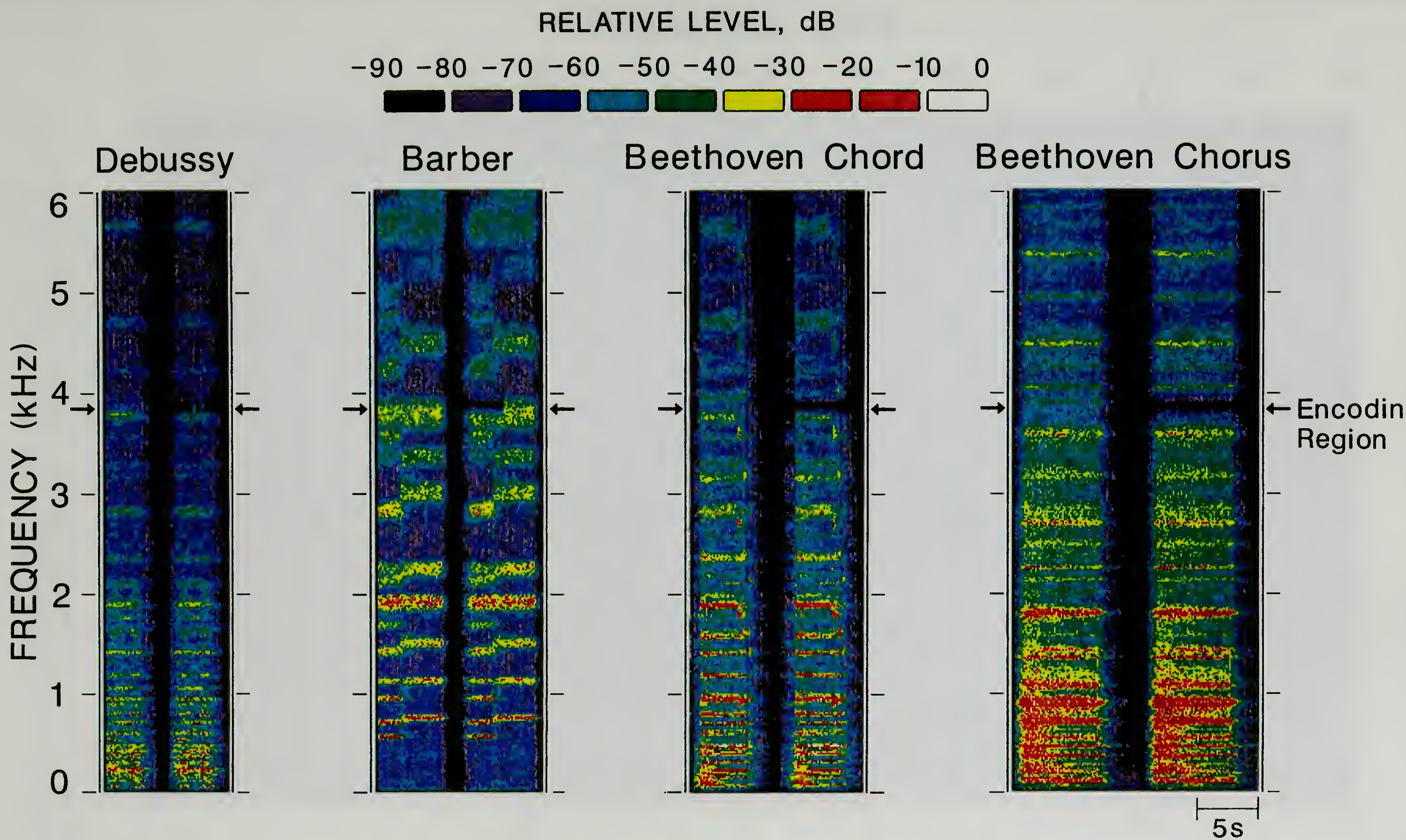


Figure 32. Color spectrographs of direct (left) and encoded (right) material, for the serial listening study, from Debussy, Barber, Beethoven chord, and Beethoven chorus selections (see Appendix E for a more complete description of these selections).

Table 6. Average score summary for all listeners in the serial listening tests. Percentage scores by selection.

Selection Name	Number of Responses	Percent Correct
Debussy	348	52.6
Barber	348	49.7
Beethoven chord	348	55.5
Beethoven chorus	348	43.7
Bizet	348	55.5
Wilson	348	50.0
Striesand	348	49.1
Beethoven Triangle	348	54.0
McGovern	348	53.2
Weather	348	46.3
Copland	348	54.9
Bernstein	348	49.1
Pink Floyd	348	48.6
S1	348	90.5
Prokofiev	348	70.1
S2	348	54.3
Metheny	348	54.0
S3	348	55.5
S4	348	49.4
Messiaen	348	56.3
Bach	348	51.4
Respighi	348	62.9
Handel	348	49.4
Ravel	348	55.7

specifically to be an easily discernible case, and the results show clearly that it was indeed easy to detect.

The second highest average score was obtained for the Prokofiev selection. In this case, 70.1% of the responses for this selection correctly distinguished whether the members of the pair presented were the "same" or "different" with respect to encoding. The statistical significance of this result may be described as follows: the probability of obtaining an average score of 70.1% or more from a set of 348 responses generated by a 50-50 chance mechanism is much less than one in a million (<< 0.000001).

To summarize, the significance probabilities for the 8 selections with scores above the central 0.95 probability interval were as follows:

Selection	Percent Correct	Significance Probability (Prob ≥ % obs'd by chance)
Beethoven chord	55.5	0.02
Bizet	55.5	0.02
S1	90.5	<< 0.000001
Prokofiev	70.1	<< 0.000001
S3	55.5	0.02
Messiaen	56.3	0.009
Respighi	62.9	< 0.000001
Ravel	55.7	0.017

The Beethoven chorus, whose score fell below the 0.95 probability interval, had a score of 43.7 percent correct, corresponding to a significance probability of 0.009. These results, taken together, show clearly that at least selections S1, Prokofiev, and Respighi showed discernible audible effects of encoding in the serial tests.

Average Scores for Listeners

Another way to view the data is to consider the percent scores obtained by each subject, averaged over the 24 selections (96 test items). As discussed above, the 24 selections differ substantially in difficulty of detecting the effect of the encoder. Some of the chosen musical selections were chosen with the expectation that audible effect of the encoder would be very subtle. For such selections, it is reasonable to expect that a typical listener would have about a 50% average probability of giving the correct response to the Same/Different questions in the serial listening test.

The effect of including selections having rather subtle or nonexistent audible effects of encoding is to compress the test scores for subjects in a closer range around 50% compared to the range of average scores for selections. This is because the average score computed for a given subject represents an average of the probability of correct discrimination, averaged across the 24 selections.

For example, if half of the selections were subtle, having a 50% chance of correct discrimination, and the other half had a 70% chance, the average score expected for listeners would be 60%, the average of 50% and 70%. The data suggest that at least half of the selections were subtle and that the range of probabilities of correct responses is greater than this illustration. As another example, suppose that 2/3 of the selections (i.e. 16 of the 24) were subtle (with 50% probability of correct discrimination), and that the remaining 8 selections had probabilities spread out from 60% to 95% as follows: 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%. Then the expected score for listeners on such a collection of selections would be 59.2%, the overall average of all 24 probabilities. These hypothetical examples illustrate that the result of constructing a test using a mixture of selections with subtle

and more noticeable effects is to reduce the average scores computed for listeners.

The data on average scores for earphone and loudspeaker subjects, respectively, are shown in tables C-3 and C-4 of Appendix C. Another presentation of these data is provided by figures 34 and 35. In these figures, the response behavior of listeners is compared for test items consisting of truly different pairs (DE or ED), and test items that were truly the same (DD or EE). Specifically, the figures show the percentage of times each subject replied "Different" when the test items were in fact different (vertical axis) plotted against the percentage of times that same subject replied "Different" when the test items were in fact the same (horizontal axis).

The schematic plot shown in figure 36 illustrates some extreme possibilities. A listening subject who never responded "Different" for either the DE and ED pairs or for DD and EE pairs would be plotted on the diagonal line at the point labeled "A". Similarly, a subject who always responded "Different" for both the DE and ED pairs and also for the DD and EE pairs would be plotted on the diagonal line at the point "B". In contrast, a subject who achieved perfect discrimination would be plotted far above the diagonal at point "C," which corresponds to responding "Different" 0% of the time for DD and EE items and 100% of the time for DE and ED items.

Subjects achieve scores that fall above the diagonal if, and only if, they respond "Different" more often for pairs that are truly different than for pairs that are the same. Clearly, subjects who achieved an average score of more than 50% correct are exactly those whose scores are plotted above the diagonal in these figures. The fact that the points in figures 34 and 35 are somewhat spread out from left to right in the plots is evidence that the subjects had a range of propensities to guess "Different" when they were not sure.

The figures reflect the fact that a majority of the subjects achieved scores somewhat above 50% correct -- in other words, they were more likely to describe the pair as "Different" when they were truly different than when they were truly the same. The dashed curves in the figures enclose the range within which scores would fall 95% of the time in the absence of an encoder effect.[1] That is, the dashed curves form a 0.95 probability interval that would apply to the case of all responses amounting to 50-50 guesses.

Several of the subjects achieved scores outside the 0.95 probability interval, so their performance can not be readily explained by "chance." Specifically, there are 10 earphone subjects and 3 loudspeaker subjects whose average scores are significantly higher than 50%, with (one-sided) significance probabilities ≤ 0.025. Among the 84 (=61+23) subjects represented in figures 34 and 35, it

---

[1]The dashed curve in the figures is constructed so that a plotted point, (x,y) is inside the curve if $|x-y| \leq 2[2p(1-p)/48]^{\frac{1}{2}}$, where $p = (x+y)/2$. This corresponds to the usual procedure for testing the hypothesis that two proportions, x and y, are both estimates of the same probability.

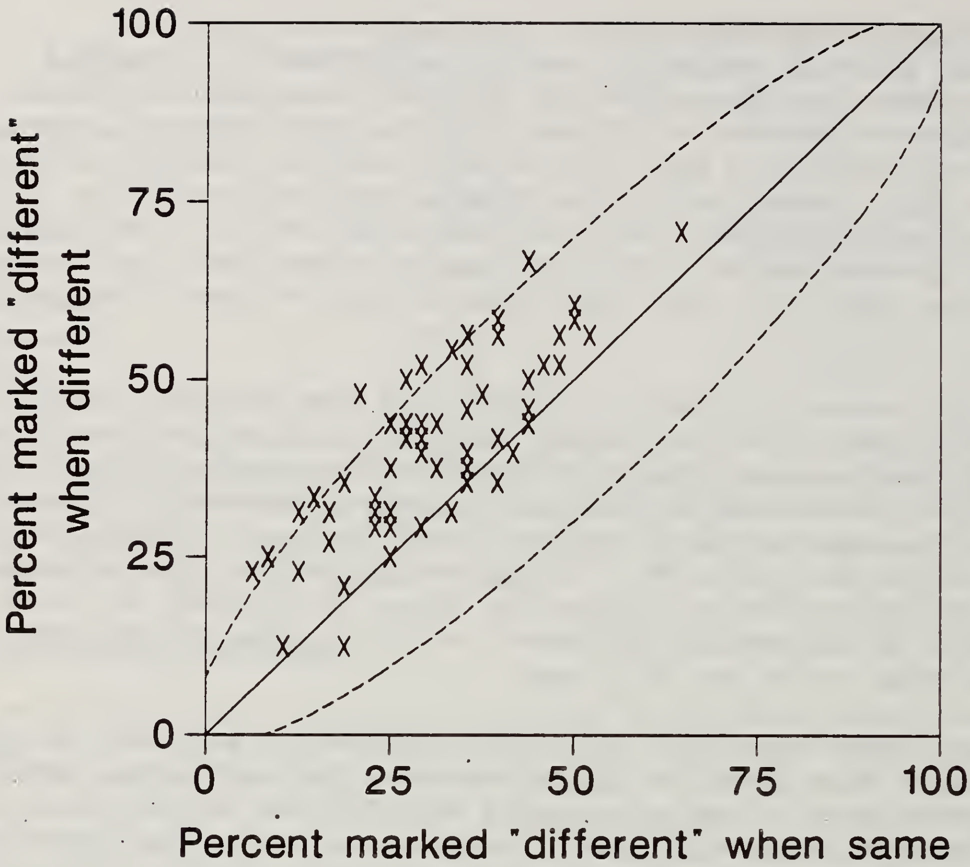


Figure 34 "Signal detection" plot for n=61 earphone listeners. Each point (x,y) represents one subject whose plotting position is defined as x = percentage of DD and EE test items to which the subject answered "Different," and y = percentage of DE and ED test items to which the subject answered "Different." Note that y is a percentage of correct answers, while x represents wrong answers. The dashed curves enclose a region in which points would fall with probability 0.95 if there were no audible effect of the encoder. (The overall percentage of correct answers for a given subject is $[y + (100-x)]/2$, which exceeds 50% when $y > x$, i.e. when the point is above the diagonal line.)

might be expected that 2 or 3 ($0.05 \times 84 = 2.1$) would lie outside the 0.95 probability interval by chance. The fact that the scores for 13 subjects fell outside the interval, with all lying above the interval, supports the conclusion that the result is not due to chance but rather to the presence of a real average audible effect for the selections in the test.

A prominent qualitative feature of the figures is the preponderance of data above the diagonal "chance" line. This reflects the fact that a substantial majority of subjects scored above 50% on the tests. We now consider whether the size of this majority could be due to chance.

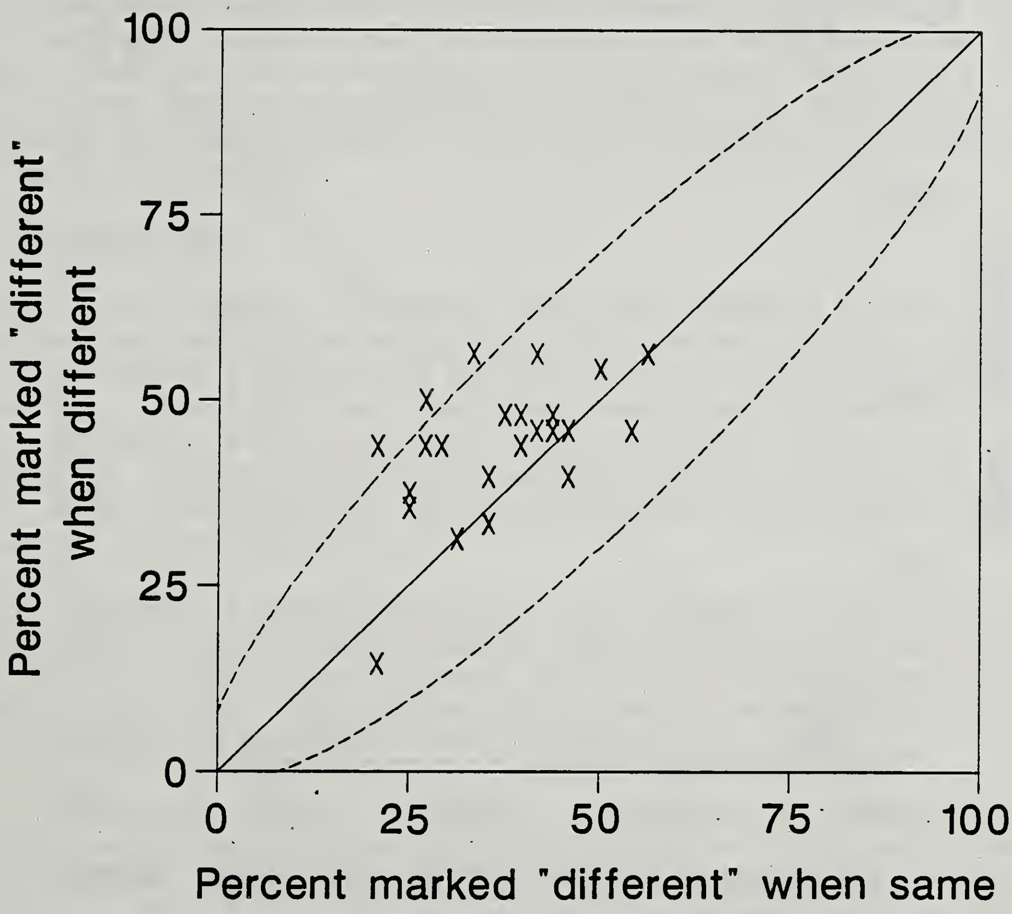


Figure 35 "Signal detection plot for n=23 loudspeaker listeners. The interpretation of this figure is the same as that for figure 35.

The statistical significance of the tendency of results to fall above the diagonal in the figures (scores above 50% in tables C-1 and C-2) can be evaluated by the sign test, as follows. Of the 84 (=61+23) subject scores represented in figures 34 and 35, 69 of them lie above the diagonal, 7 are exactly on the diagonal (for subjects who scored exactly 50% correct), and 8 are below the diagonal. Under the null hypothesis of no audible encoder effect, the probability of achieving a score above 50% would be 0.5. In that case, a calculation using the binomial probability distribution shows that the probability of a obtaining 69 or more scores above 50% (and 8 or less below 50%) is about 3×10^{-12}, or 3 in one trillion. Thus the data strongly suggest, through the persistent tendency to achieve scores above 50%, that there is an audible encoder effect reflected in the subjects' responses to the test selections.

It may be recalled that the 24 music selections chosen for the test include 4 selections that were created using a keyboard synthesizer. To investigate the degree to which the strongly significant result of the sign test above is due to the presence of synthesizer selections in the serial test, we computed the average scores per subject after deleting the synthesizer selections. In this case the average scores were based on 80 test items per subject, rather than 96 for the

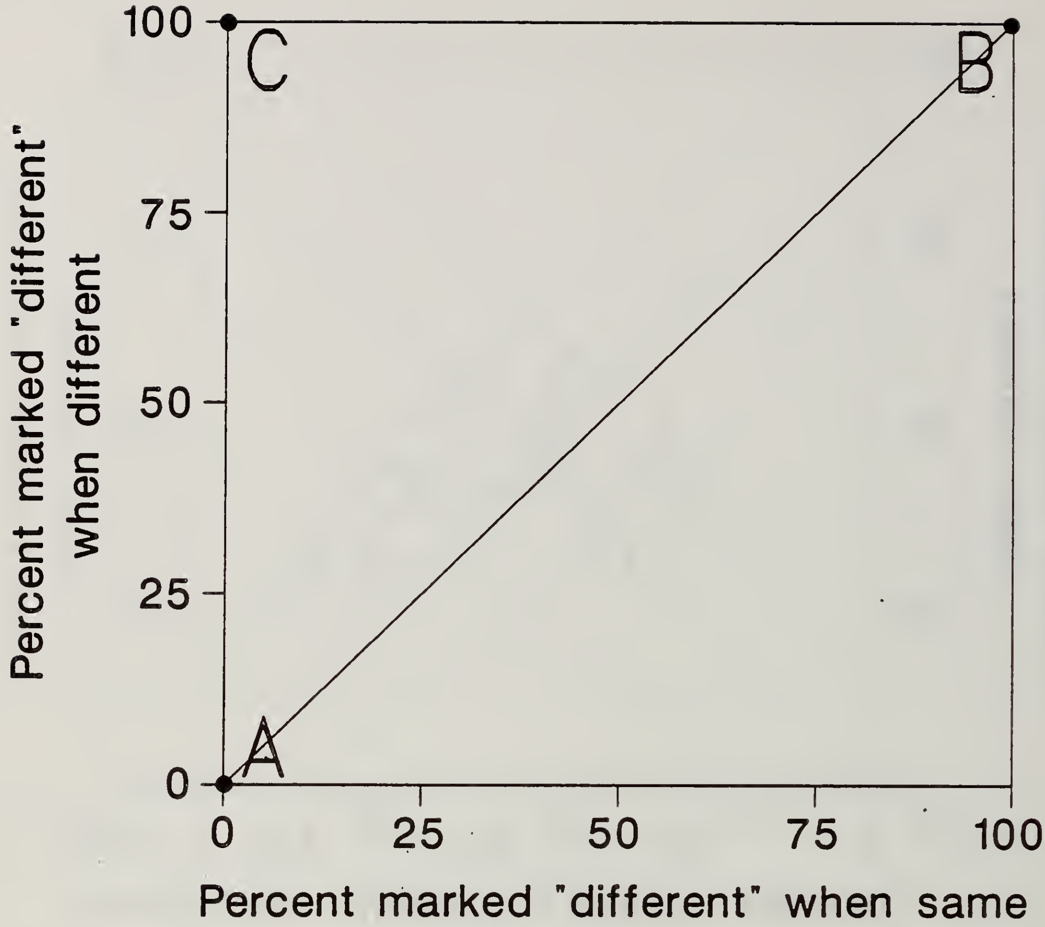


Figure 36 Schematic plot of responses for listening subjects. Point A corresponds to a listener who answers "Same" for all test items, whether they are truly same or different. Point B corresponds to a listener who always answers "Different." Listeners who answer "Different" the same percent of the time to both truly same and truly different pairs will fall on the diagonal line. Point C corresponds to a listener who discriminates perfectly between same and different pairs.

full test. For the reduced test data, consisting only of recorded music selections, a total of 58 of the 84 earphone and speaker subjects achieved scores greater than 50%. The number of subjects with scores of exactly 50% was 4 and the number with scores below 50% was 22. For the total of 84 subjects, the sign test evaluates the probability of getting 58 or more scores over 50% (and 22 or less below 50%) assuming chance responses. The calculated probability is 0.000035 in this case, so the observed excess of scores above 50%, compared to the number below 50%, would happen only about 3.5 times in 100,000 assuming chance responses. Thus, while the significance probability is somewhat less extreme than was obtained for the full test of all 96 items, the reduced test data still show quite strong evidence of an audible effect.

This analysis of average scores for listeners was, for convenience, based only on the 84 subjects whose scores are plotted in figures 34 and 35, so the three listeners who used a mixed format of earphones and loudspeakers were not included. Since these three additional subjects all scored higher than 50% on both the full test and also the reduced test obtained by dropping the synthesizer selections, their inclusion in the calculations would only strengthen the evidence for an audible effect of encoding.

Parallel Listening Test

The results of the parallel listening tests are presented in table 7. Again, the data show that the ability to detect an audible encoder effect varies substantially between subjects and between musical selections. Some of the selections (e.g., Respighi) that were most consistently detected in the serial tests were apparently harder to detect in the parallel format tests. Some selections (e.g., Metheny) were detected more often in the parallel format than the serial tests. The Prokofiev selection achieved consistently high scores in both the parallel and serial tests.

To evaluate statistical significance for the listeners' scores, it can be shown that the probability that a given subject would be able to correctly identify 9 or 10 out of 10 selections, assuming there is no audible effect and all responses are 50-50 guesses, is 0.011. So, subjects D, G, and O scored significantly higher than can reasonably be accounted for by chance ($p \leq 0.011$).

Turning to the average scores for musical selections, the probability that any given musical selection would be correctly identified 12 or more times out of 15, assuming no audible effect, is 0.018. Since the encoded track was correctly identified by 12 out of 15 subjects for both the Metheny and the Prokofiev selections, these results are statistically significant at the $p = 0.018$ level. (Among the four selections that had only 14 responses, the corresponding probability of getting 10 or more correct, the highest score observed, is 0.09.)

Conducting independent significance tests repeatedly for 24 selections could be expected, on average, to produce about 0.4 significances at the 0.018 level ($0.4 = 0.018 \times 24$). The observed data produced results at the 0.018 significance level for two selections, and in addition two other selections achieved scores of 11 out of 15 correct (significance probability of 0.059), amounting to substantially more significances than would be expected by chance. Together, these calculations show that the test results are not well-explained by chance, even after allowing for the fact that multiple significance tests were performed.

In summary, the data for the Metheny and Prokofiev selections, at least, show statistically significant evidence ($p < 0.02$) that the encoded channel is audibly distinguishable from the unencoded version. For the remaining selections, the proportion of times they were correctly identified, although quite high in some cases, is not based on a large enough sample of subjects to show a statistically significant difference (at the $p < 0.05$ level) from 50%.

Table 7. Subject responses for all listeners in the parallel listening tests

Selection	Subject															Fraction Correct
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	
1. Bizet	C	•	C	C	C	w	C	C	w	w	w	w	C	C	w	8/14
2. Streisand	w	C	C	C	w	C	C	w	C	C	C	C	w	C	C	11/15
3. Beethoven Triangle	C	•	C	C	C	w	C	w	C	C	w	C	w	C	C	10/14
4. Copland	w	w	C	C	C	w	C	w	C	w	w	C	w	w	C	7/15
5. Metheny	C	C	C	C	C	C	C	C	C	C	w	w	C	w	C	12/15
6. Prokofiev	C	w	w	C	C	C	C	C	C	C	w	C	C	C	C	12/15
7. Bernstein	C	w	C	C	w	C	C	w	w	C	C	C	C	C	C	11/15
8. Messiaen	C	•	C	C	w	C	w	w	C	w	w	w	w	C	C	7/14
9. Bach	w	•	C	C	C	w	C	w	C	w	w	C	C	w	C	8/14
10. Respighi	w	w	w	C	C	w	C	w	w	w	w	w	w	C	C	5/15
Total Correct	6	2	8	10	7	5	9	3	7	5	2	6	5	7	9	

C = Correct

w = Wrong

◦ = No Response

Average percent correct: $100 \times (91/146) = 62.3$

5. CAN THE SYSTEM BE BYPASSED, AND IF SO, HOW EASILY?

5.1 Interpretation of the Question

This question is primarily one dealing with the Decoder portion of a DAT Recorder/Decoder. Since the CBS-supplied DAT Recorder/Decoder is a prototype unit, access to individual components may not be representative of the final design. Therefore, only the possibilities of circumventing the Decoder by means that do not require physical modification methods were considered, along with their relative ease of implementation.

5.2 Description of Defeat Methods

There are basically two external signal conditioning methods of circumventing the Decoder operation. One approach is to add electronically to the incoming signal of the DAT Recorder an arbitrary signal with frequency components in the notch region that serve to prevent a POSSIBILITY activated and RECORD INHIBIT state. The total signal would then be recorded on the DAT Recorder/Decoder, but could likely be audibly objectionable. Upon playback of such a digital audio tape, however, the added "defeat" signal could be subtracted electronically from the recorded signal to restore the resultant audio signal to nearly what was input to the DAT Recorder/Decoder. That is, the listener of the resultant signal would hear the original signal (as encoded), plus any residual, uncompensated components of the "defeat" signal. Because of the need for additional playback circuits and the lack of a digital audio tape which can be copied and played without the need for special playback circuits, NBS did not consider this approach during the evaluation of the CBS copy prevention code system.

The second (and more preferable) external approach is to add electronically to the incoming signal of the DAT Recorder/Decoder just enough of a signal (with frequency components in the notch region) to compensate for what the encoding (notch filtering) process removes from the original signal. Ideally, the spectrum (both amplitude and phase) of the original audio signal would be perfectly reconstructed before being input to the DAT Recorder/Decoder. Of course, perfect reconstruction is not possible (due to noise, imperfect electronic components, etc.), even if the encoder's notch filter characteristics are known in great detail. However, several schemes for filling in the notch region amplitude spectrum have been attempted during this evaluation, and are described below.

<u>Defeat Method 1</u>

Figure 37 shows a basic means for adding to the encoded input signal to be recorded another signal, such that the amplitude spectrum of the output signal no longer contains enough of a notch in the spectrum to cause a RECORD INHIBIT state to occur. The added signal is derived from a wideband noise source with its output passed through a bandpass filter, centered in the notch region at about 3.8 kHz. By adjusting the amplitude of the added signal, the notch in the input signal spectrum is effectively filled in. This method is referred to as Method 1. There are numerous possible electronic circuit

implementations of this method. Shown in Appendix I is the schematic diagram of a noise source, a 3.8 kHz bandpass filter, and a summing amplifier constructed at NBS for purposes of implementing Method 1.

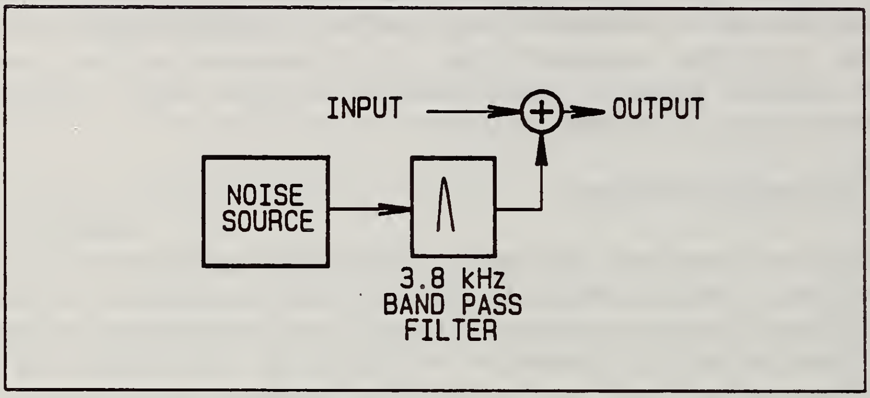


Figure 37 Defeat method 1 that uses a bandpass-filtered noise source added to the encoded input signal.

Figure 38 displays the spectrum obtained for an example of an input signal from a broadband noise source which has been encoded with an obvious notch in its spectrum (shown in green). This signal, of course, caused the RECORD INHIBIT state to occur in the Decoder. The blue spectrum shown is that of the added bandpass-filtered noise source (Method 1, as described above and with the circuits given in Appendix I). The resultant output signal spectrum is shown in red (offset for clarity), and can be seen no longer to contain a notch in its spectrum. This signal, of course, did not cause RECORD INHIBIT in the Decoder.

Defeat Method 2

Alternatively, the source for a defeat signal can be derived from the encoded input signal itself. Figure 39 shows one method for obtaining such a defeat signal, which is obtained by passing the input signal through a bandpass filter an octave below the notch band, full wave rectifying the 1.9 kHz bandpass filter output (to double the frequency), and then filtering the rectifier output with a 3.8 kHz bandpass filter. This defeat signal is then added to the input signal as shown, and the resultant output (sent to the DAT Recorder/Decoder) will have its amplitude spectrum continuously adjusted in the notch region by the amount of energy contained in the original audio signal at around an octave below the notch center frequency. This method is referred to as Method 2. The assumption in this method is that if there was an encoding notch placed in the incoming signal, then there is a likelihood that the original, unencoded audio signal had spectral components at an octave below the notch. Detecting the presence of these components in the input signal can determine when (and in what amount) to add a defeat signal at the notch band to the input signal. Shown in Appendix I is the schematic diagram of a 1.9 kHz bandpass filter, a full wave rectifier, and the other components constructed at NBS for purposes of implementing Method 2.

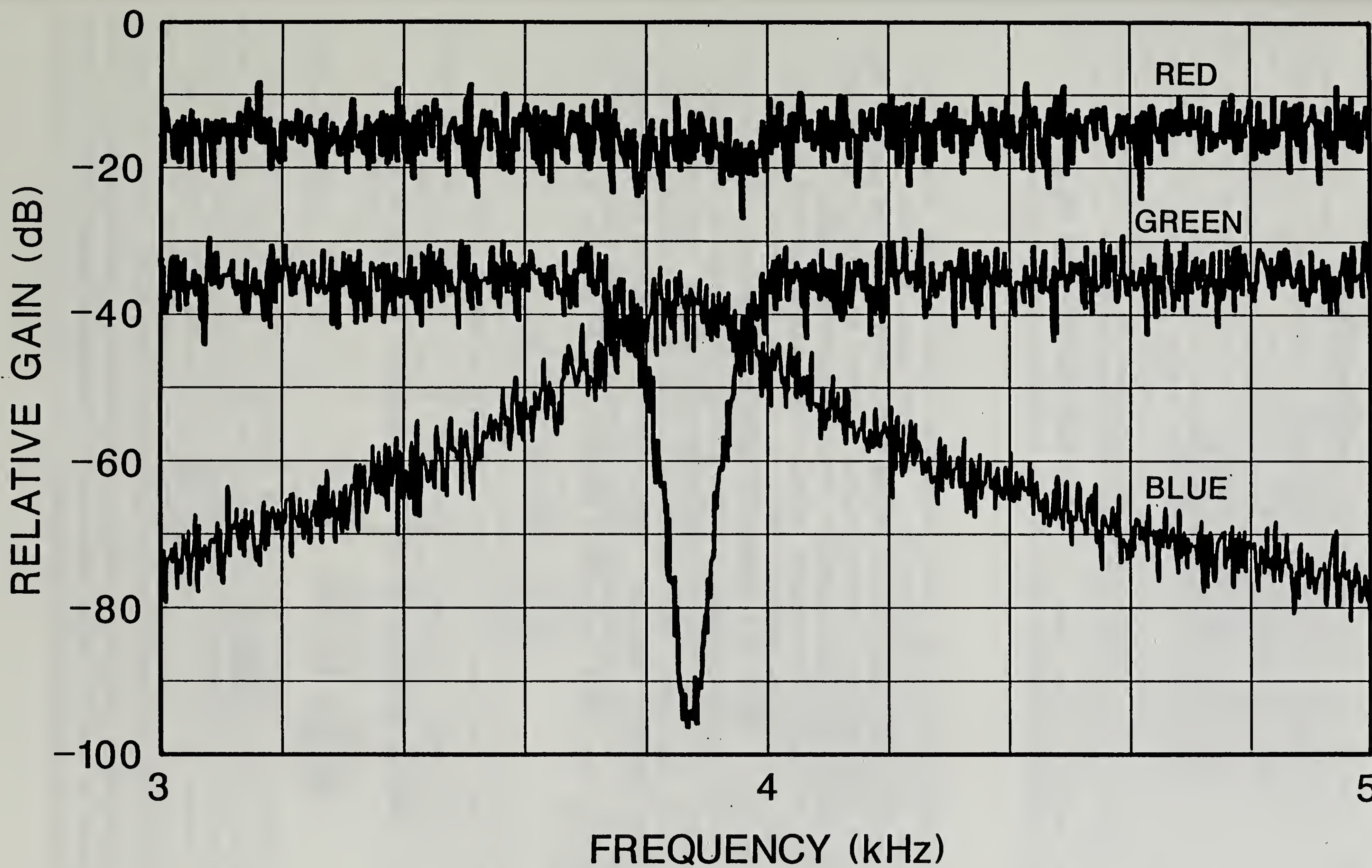


Figure 38 Frequency spectra of: (1) a broadband noise input signal containing the encoding notch (green spectrum), (2) the added bandpass filtered noise source (blue spectrum), and (3) the resultant output signal (red spectrum).

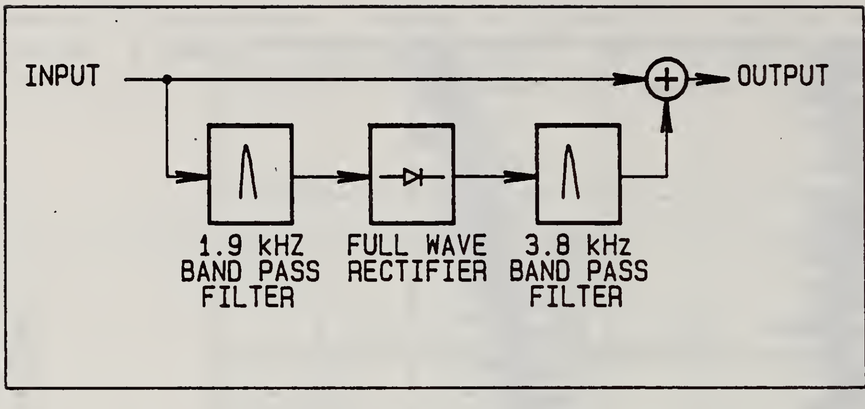


Figure 39 Defeat method 2 that uses a 1.9 kHz component of the input signal doubled to match the 3.8 kHz encoding notch frequency.

Defeat Method 3

Figure 40 shows a variation of the basic bandpass-filtered noise source used as the added defeat signal (see figure 37). In this case, however, rather than continuously adding the bandpass-filtered noise source to the input, this signal is amplitude modulated by another signal derived from the input signal before being summed with the input signal. A dc modulation signal, proportional to the frequency components of the input signal around 2.7 kHz, is obtained with the use of the bandpass filter, full wave rectifier, and lowpass filter circuits. Thus, the amount of bandpass-filtered noise (at 3.8 kHz) added to the input signal depends on the amount of energy in the input signal spectrum around 2.7 kHz. With the presence of signal components in this part of the input spectrum, the defeating signal in the 3.8 kHz region is then added in a more "masked" fashion. This method is referred to as Method 3. Shown in Appendix I is the schematic diagram of a 2.7 kHz bandpass filter, a lowpass filter, and the other components constructed at NBS for purposes of implementing Method 3.

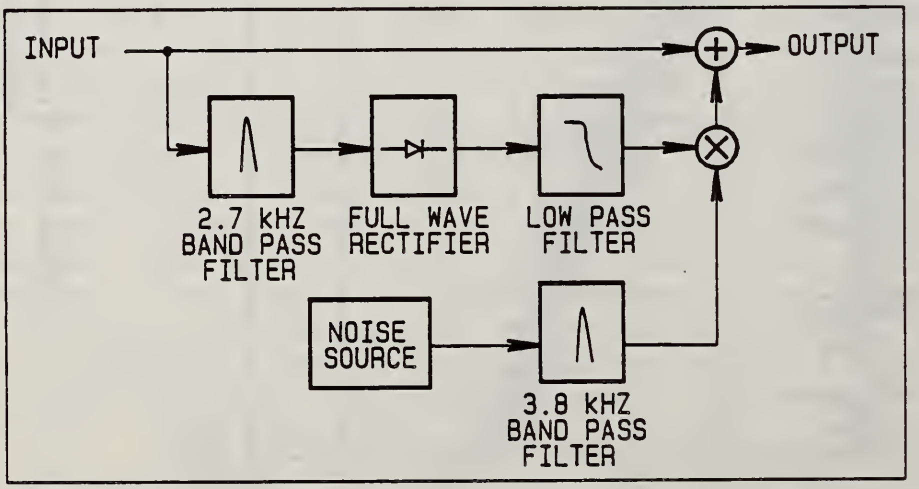


Figure 40 Defeat method 3, which is similar to method 1, and uses the amplitude of the 2.7 kHz component of the input signal to modulate the bandpass-filtered noise.

Defeat Method 4

Figure 41 is very similar to the circuit of figure 40, except that the source of the defeat signal, rather than being bandpass-filtered noise, is derived from the input signal, as discussed previously. Thus, not only is the amount of added defeat signal dependent on the frequency components in the input signal at an octave below the notch, this signal is also amplitude modulated by a dc modulation signal proportional to the amplitude of the frequency components of the input signal around 2.7 kHz. The encoding notch of the input signal is thereby filled in dynamically by the amount of energy at two lower frequency bands of the original audio signal. This method is referred to as Method 4. Shown in Appendix I are all the components constructed at NBS for purposes of implementing Method 4.

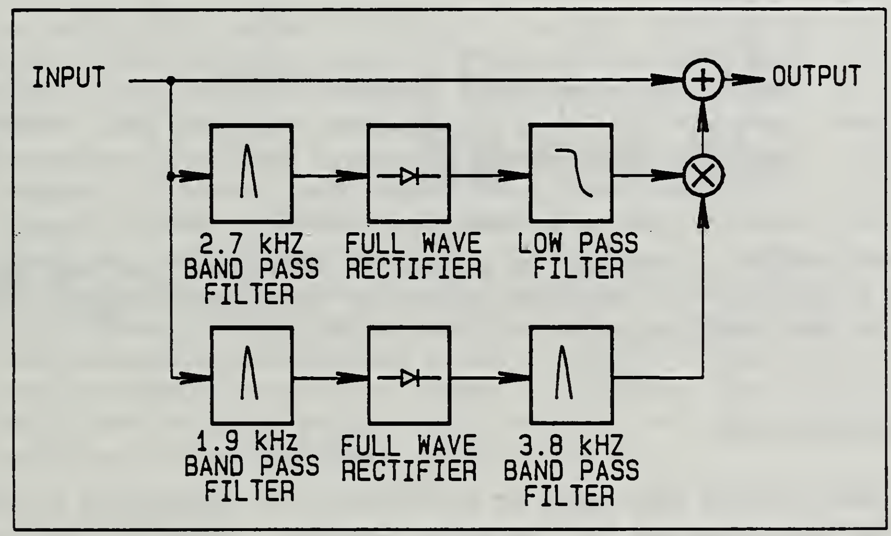


Figure 41 Defeat method 4, which is similar to method 2, and uses the amplitude of the 2.7 kHz component of the input signal to modulate the frequency doubled 1.9 kHz component of the input signal.

Defeat Method 5

A further variation of this dynamic modulation theme is provided in figure 42 where the dc modulation signal of figure 41 is derived from the amount of input signal energy in either of the two notch sidebands, centered at about 3.5 and 4.3 kHz. The assumption in this case is that the amount of added defeat signal (besides being dependent on having input signal components at an octave below the notch band) should be proportional to the amplitude of the input signal spectrum on either side of the notch band. If there are no appreciable sideband components, then there is not likely to be either a purposeful (encoding) notch or a "natural" notch in the amplitude spectrum of

the incoming signal. Of course, this modulation scheme could also be applied to a bandpass-filtered noise source as well. This method is referred to as Method 5. Shown in Appendix I is the schematic diagram for a 3.5 kHz bandpass filter, a 4.3 kHz bandpass filter, and all the other components constructed at NBS for purposes of implementing Method 5.

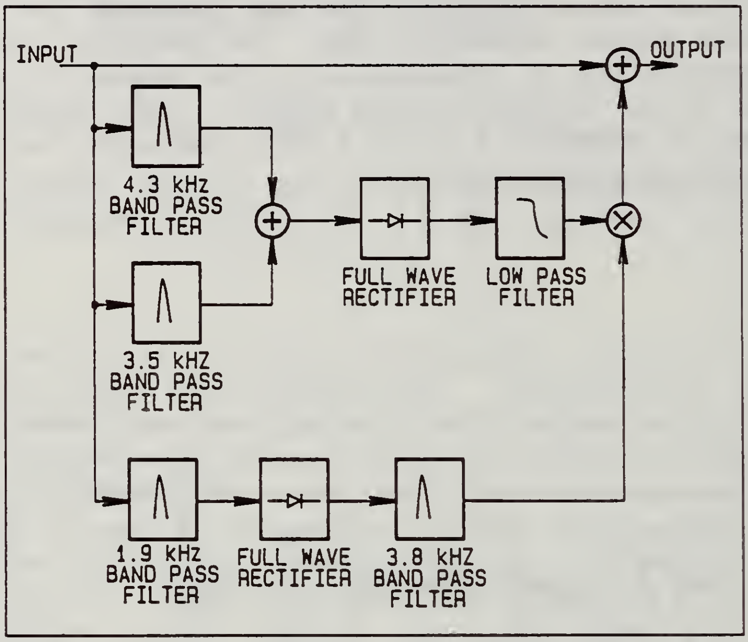


Figure 42 Defeat method 5, similar to method 4, and uses the amplitude of the 3.5 kHz and 4.3 kHz components of the input signal to provide the modulation.

5.3 Ease of Implementation

In an effort to demonstrate the ease or difficulty of bypassing or defeating the Decoder in the CBS-supplied DAT Recorder/Decoder, several electronic circuits have been assembled that bypass the Decoder, which are based on the defeat methods described in section 5.2 above. As shown in Appendix I, the ease of bypassing can be judged by the relative simplicity involved in these circuits, and how easy they are to implement.

The basic active electronic components in all of these circuits are semiconductor diodes and operational amplifiers. Together with passive resistor and capacitor elements, most of these circuits can be implemented with about a dozen commonly available, off-the-shelf parts. The power supplies needed to operate these circuits are also available as standard modules. It is estimated that the parts cost in small quantities of the circuits required for each of the methods described in this report is less than $100, including the dc power supplies.

The degree of knowledge required to conceive, design, and test the defeat circuits shown in this report is generally that of a professionally trained

electrical/electronics engineer. However, the degree of knowledge required to construct the defeat circuits described herein is that of an electronics technician, familiar with reading circuit schematics. For the very simple method of adding a signal from an oscillator (at or near the notch frequency) to the encoded audio signal, in order to defeat the Decoder in a DAT recorder, this approach and degree of knowledge is available to anyone who reads about the subject of the proposed CBS copy prevention protection scheme described in popular audio magazines. On the other hand, to determine just how well a given defeat method can be made to work requires test equipment, such as a spectrum analyzer, and detailed knowledge of how the Decoder circuit operates, knowledge generally not available to a novice.

5.4 Audibility of Defeat Methods

All of the circuits shown in Appendix I were constructed for implementing the five defeat methods described in section 5.3 above. The relative audible effects of the defeat methods were informally evaluated by a subjective listening test made by NBS staff. Table 8 gives a summary of the results of this test.

The relative audibility shown in table 8 for the five defeat methods is that observed by the NBS listening staff. No claim is made that this table is representative of the audible effects of these five methods on an absolute basis. However, the relative audibility that the NBS staff observed does appear to correlate with the degree of circuit "complexity" associated with the defeat method. However, as can be seen from the circuits shown in Appendix I, none of the circuits implemented at NBS for the five methods described above are considered complex. Continuously adding bandpass-filtered noise at the notch band (Method 1), for example, is a relatively simple method, but it causes a noticeable audible hiss, particularly during quiet passages. The circuits for Methods 2 and 3 are of intermediate relative complexity and produced occasional audible effects. The defeat circuits for Methods 4 and 5 are relatively the most complex, and these provided relatively no audible effects in the informal tests.

TABLE 8

Audibility of Defeat Methods

Defeat Method	Relative Audibility
1. "Noise"	Noticeable
2. "1.9 kHz"	Occasional
3. "Noise - 2.7 kHz Mod"	Occasional
4. "1.9 kHz - 2.7 kHz Mod	Inaudible
5. "1.9 kHz - 3.5/4.3 kHz Mod"	Inaudible

6. SUMMARY OF RESULTS

6.1 Does the System Achieve Its Purpose?

As stated in section 3.1 under the interpretation of this question, the copy protection code system, as embodied by the CBS-supplied Encoder and DAT Recorder/Decoder will achieve its purpose if the Decoder inhibits the CBS-supplied DAT Recorder/Decoder from recording all or parts of an encoded signal and does not inhibit the recording of unencoded signals. A signal is considered encoded if it has passed through the Encoder; otherwise, it is unencoded.

The evaluation of the two CBS-supplied Encoders has been described in section 2.2 (Physical Measurements) and section 4.2 (Effect on the Performance Specifications of an Audio System). The evaluation of the performance of the Decoder portion of the CBS-supplied DAT Recorder/Decoder has been described in section 2.4 (Physical Measurements), and section 3. (Does the System Achieve Its Purpose?). As summarized below, the answer to this question on the basis of these results is no - the system does not achieve its purpose.

<u>Encoder Performance</u>

The Encoder uses a narrow-band filter to "notch" the music at a nominal center frequency of 3840 Hz, approximately midway between the highest B-flat and B of a piano. This encoding notch filter is in excess of 80 dB deep near its center frequency and has a bandwidth of nominally 220 Hz. The phase characteristic of the notch filter varies drastically over the bandwidth of the filter. The Encoder is dynamic, in that it starts and stops notching, depending upon the signal components near 3840 Hz and near 2715 Hz. The two Encoders that were supplied by CBS Records differ considerably as regards the switching in and out of the encoding notch filter. One of the Encoders frequently switches the notch filter out -- this results in there being no encoding much of the time, resulting in many false negatives (see below). The other encoder leaves the notch filter in much more of the time, potentially resulting in increased audibility of the removal of part of the input signal.

<u>Decoder Performance</u>

The primary reason that the system does not achieve its purpose has to do with the performance of the Decoder. The physical measurements made on this Decoder have shown that it is subject to both inhibiting the recording of unencoded signals (false positive detection) as well as not inhibiting the recording of encoded signals (false negative detection).

The tests that are described in section 2.4 (under Actual Encoded and Unencoded Audio Signals input to the Decoder) were certainly definitive in terms of demonstrating that there are numerous cases where the Decoder failed to detect legitimate, encoded signals from its own Encoders (false negatives) and, more importantly, failed to accept unencoded input signals of several varieties (false positives). Although these tests were not exhaustive, because of the practically infinite number and kinds of possible source material, it was observed that out of 54 different selections of unencoded

compact discs (CDs), played on a standard CD player and input to the DAT Recorder/Decoder, 10 of these CDs had tracks and sectors that consistently caused a RECORD INHIBIT state (false positive detection), and shut down the DAT Recorder. These 54 CDs contained 502 tracks of a wide variety of musical selections. False positive detection was observed on 16 of these 502 tracks. Especially significant is the fact that organ and violin recordings consistently produced false positive detection. Similarly, in tests with encoded signals, the Decoder did not cause a RECORD INHIBIT (false negative detection) at least once for 8 out of 18 CDs encoded using Encoder 006 nor for 3 out of 32 CDs encoded using Encoder 012. On a per track basis, the Decoder made a false negative decision on 139 out of 150 tracks encoded using Encoder 006 (~93 percent false negative detection rate) and on 112 out of 280 tracks encoded using Encoder 012 (~40 percent false negative detection rate).

Because of the scanning operation of the Decoder, described in sections 2.3 and 2.4, there is a variation associated with the false positive detection of the Decoder. Consequently, as described in section 3.2, the number of false positives observed during the tests for these occurrences varied. Depending on the length of the time used for monitoring the input signal, and the number of repeat monitorings, the total number of false positive detections varied. For the NBS tests, the percentage of false positive detections ranged from 8 to 86 percent.

The other tests that are described in section 2.4 under Physical Measurements have also helped to substantiate that the Decoder in the CBS-supplied DAT Recorder/Decoder does not perform well in discriminating a purposeful encoding notch in the amplitude spectrum of an input signal. Using a combination (three-tone) sine wave input signal (with frequency components in the notch region), it was determined that the measured Decoder response characteristics are very dependent on both the absolute and relative amplitudes of the notch band and the upper and lower sideband levels of the input signal. These relationships are shown not to require a significant notch type of characteristic in the notch frequency region of the input signal spectrum in order to cause a RECORD INHIBIT state under certain conditions. Other input signal levels at the notch and upper and lower sideband frequencies may require a very deep notch characteristic in the input signal in order to cause a RECORD INHIBIT state. This limited discrimination characteristic is further revealed by the NBS tests using functional triangle, sawtooth (ramp), and square wave input signal waveforms. It has been determined that there is a range of amplitudes and (fundamental) frequencies of these waveforms which satisfies the spectral requirements needed to cause the POSSIBILITY and RECORD INHIBIT states in the CBS-supplied DAT Recorder/Decoder. However, these amplitude and frequency combinations were not explored exhaustively.

6.2 Does the System Diminish the Quality of the Prerecorded Material?

If the Encoder were to be introduced into the music recording and reproduction process, the frequency characteristics of the signal coming out of an amplifier in one's home would be degraded relative to the signal that was originally recorded prior to copy-prevention coding. Specifically, the usual statements of system performance as being within ±X dB (amplitude) and ±Y deg (phase) would no longer be met.

In order to examine directly the audibility of the effects of the Encoder, two separate but related series of listening tests were carried out. In each of these tests, listeners attempted to discriminate between music which was presented to them "directly" (not encoded) and the same musical selection presented to them "encoded." In the "serial" listening tests, 24 short (of the order of 5 sec) segments of music were presented sequentially to multiple listeners in a fixed sequence that was established by NBS staff. Pairs consisting of direct (D) and encoded (E) copies of each segment were presented to the listeners, with all possible combinations, DD, DE, ED, and EE, included for each segment in the test. The listeners were asked to identify whether the copies in each pair presented were the same or different.

In the "parallel" listening tests, ten segments of music, each of about 1 minute in length, were presented to one subject at a time. The selections were recorded synchronously on three parallel pairs of tracks, with one reference track-pair always being direct material, a second track-pair being direct or encoded material, and the third track-pair being encoded or direct. The particular track-pair containing encoded material was assigned randomly. The subject's task was to attempt to identify, for each musical selection, the track-pair containing the encoded material.

Generally, the results of the subjective listening tests show that the ability to hear effects of the Encoder varies substantially among individual subjects and, especially, among musical selections.

Of the 24 musical selections in the serial listening study, eight of these have average scores that exceed 50 percent (i.e., pure chance) by a statistically significant margin ($p < 0.025$). For three of these selections, the data showed evidence of an audible Encoder effect that was strongly statistically significant, with the probability of obtaining such high scores by chance being less than one in a million. Several listeners achieved scores that exceed chance by a statistically significant margin ($p < 0.025$). More significant, statistically, was the fact that 69 of 84 listeners achieved scores higher than 50 percent on the test, a result that would occur by chance with a probability of only 3 in a trillion.

In the parallel listening study, three listeners correctly identified the encoded channel for at least 9 of the 10 selections, a result that is statistically significant at the 1 percent level. For two selections, the encoded track-pair was correctly identified often enough to achieve statistical significance ($p < 0.02$). Of these selections, one was the same as the highest scoring compact disc selection in the serial listening tests, while the other did not achieve a statistically significant score in the serial tests. This difference suggests that a longer-term phenomenon may have been detected in the parallel study (which used a 29-second segment) that was not discerned in the serial test (which used a 5-second segment).

6.3 Can the System be Bypassed, and If So, How Easily?

Several possible methods by which the Decoder used in the CBS-supplied DAT Recorder/Decoder can be circumvented or defeated by external signal

conditioning methods have been demonstrated. Five such methods are described in section 5.2 of this report. Electronic circuits for all of these methods were implemented during the NBS study. Besides the functional description of these methods, the schematic diagrams of the components used in the circuits implemented at NBS are provided in Appendix I of this report. There are certainly other possible methods besides the ones described herein for externally circumventing the Decoder.

As shown in Appendix I, the ease of bypassing can be judged by the relative simplicity involved in these circuits. The electronic components in all of these circuits are very basic. Most of these circuits can be implemented with about a dozen commonly available, off-the-shelf parts. The power supplies needed to operate these circuits are also available as standard modules. It is estimated that the parts cost in small quantities for each of the circuits shown in Appendix I is approximately $100, including the dc power supplies. The degree of knowledge required to conceive, design, and test the defeat circuits shown in this report is generally that of a professionally trained electrical/electronics engineer. However, the degree of knowledge required to construct the defeat circuits described herein is that of an electronics technician, familiar with reading circuit schematics.

To be an acceptable defeat method, the effects of the method must also not be perceptibly audible. The relative audible effects of the five defeat methods implemented at NBS were compared informally by subjective listening tests made by NBS staff. Two of the five defeat methods were found to provide practically no audible effects.

Since the question of whether the system can be defeated is one dealing with the Decoder portion of a DAT Recorder/Decoder, and external signal conditioning methods have been found which are practically inaudible, the answer to this question is yes - the system can be effectively circumvented with relative simple external electronic circuitry.

Acknowledgements

The authors gratefully acknowledge the many contributions made by others in the preparation of this report. Without their help a document such as this one would not have been possible. It certainly can be said that the results obtained in this study have been the product of a team effort.

Non-NBS

Messrs. Peter Dare and Gus Skinas, Sony Corporation of America, for the loan of professional digital audio recording equipment and technical assistance in its use.

Mr. Almon H. Clegg, Denon America, Inc. and Mr. T. Shirai, Nippon Columbia Co., Ltd., for providing technical information regarding the Denon DCD-3300 compact disc player.

Mr. Farrell Becker, Audio Artistry, for consultation regarding the acoustical treatment of the loudspeaker listening room.

Dr. Floyd E. Toole, National Research Council of Canada, for helpful details regarding loudspeaker listening evaluations and measurements.

Ms. Lauri Jackson, Hearing Assessment Center, for determining hearing threshold levels of the subjects in the serial listening study.

Mr. Allen Alexopulos, Hewlett-Packard Co., for provision of the software used in the preparation of real-time color spectrographs.

NBS

Messrs. Nile M. Oldham, Donald R. Flach, Mark E. Parker, and Robert H. Palm, for critical calibrations of instrumentation and preparation of figures and graphs.

Mr. Franklin R. Breckenridge, for technical assistance in analysis and testing, for help in selection of musical passages for the listening tests, and for assistance in the acoustical treatment of the loudspeaker listening room.

Ms. Susannah Schiller, for help with statistical computing and programming.

Messrs. Raymond J. Mele and F. Keith Mackley of the NEL Graphics group, for their help in the final layout of the camera-ready copy for this report.

Mmes. Judith L. Barnard, Gail E. Trupo, and Sandra L. Carlsen, for their help in typing the numerous drafts of this report.

Messrs. J. Franklin Mayo-Wells, George A. Sinnott, and Samuel Kramer for critical and timely reviews of drafts of this report.

Finally, the authors would especially like to thank John W. Lyons, Director of the NEL, for his contributions to the writing of this report, as well as his encouragement and support to the successful completion of this project.

APPENDIX A

Congressional Committee Letters

U.S. House of Representatives
Committee on the Judiciary
Washington, DC 20515-6216
Telephone: 202-225-3951

May 21, 1987

Dr. Ernest Ambler
Director
National Bureau of Standards
Gaithersburg, MD 20899

Dear Dr. Ambler:

Our subcommittees have been considering the many issues surrounding the advent of digital audio taping. As you may already know, one of those issues involves a proposal to require that a "copy-code scanner" be inserted into digital audio tape recorders, and a corresponding "notch" into prerecorded material, to prevent the machines from copying. A number of technical questions about this blocking system have arisen, and the various interested parties have agreed that the system should be submitted to an impartial entity for testing.

We therefore request that the National Bureau of Standards perform this test, which would focus on three basic questions. First, does the copy-code scanner system achieve its purpose to prevent digital audio tape machines from recording? Second, does the system diminish the quality of the prerecorded material into which the notch is inserted? Third, can the system be bypassed, and if so, how easily?

Counsel to our subcommittees will be available to provide you with further information, and to discuss the details of this test. We also expect the interested parties to submit their suggestions about how the test can be fairly and expeditiously conducted.

Dr. Ernest Ambler
Page two
May 21, 1987

Please contact Virginia Sloan, counsel to the House Subcommittee on Courts, Civil Liberties and the Administration of Justice, at 225-3926 to begin the testing process.

Sincerely,

Bob Kastenmeier

ROBERT W. KASTENMEIER
Chairman
House Subcommittee on Courts, Civil Liberties and the Administration of Justice

Dennis De Concini

DENNIS DE CONCINI
Chairman
Senate Subcommittee on Patents, Copyrights, and Trademarks

RWK:gss

ONE HUNDREDTH CONGRESS

JAMES J. FLORIO, NEW JERSEY, CHAIRMAN

JAMES H. SCHEUER, NEW YORK
JIM BATES, CALIFORNIA
HENRY A. WAXMAN, CALIFORNIA
PHILIP R. SHARP, INDIANA
CARDISS COLLINS, ILLINOIS
DENNIS E. ECKART, OHIO
BILL RICHARDSON, NEW MEXICO
JOHN D. DINGELL, MICHIGAN
(EX OFFICIO)

WILLIAM E. DANNEMEYER, CALIFORNIA
MATTHEW J. RINALDO, NEW JERSEY
DON RITTER, PENNSYLVANIA
HOWARD C. NIELSON, UTAH
JOE BARTON, TEXAS
NORMAN F. LENT, NEW YORK
(EX OFFICIO)

U.S. House of Representatives

Committee on Energy and Commerce

SUBCOMMITTEE ON COMMERCE, CONSUMER PROTECTION, AND COMPETITIVENESS

Washington, DC 20515

[illegible]
HOUSE OFFICE BUILDING ANNEX II
PHONE (202) 226-3160

GREGORY E. LAWLER
CHIEF COUNSEL AND STAFF DIRECTOR

June 3, 1987

Dr. Ernest Ambler
Director
National Bureau of Standards
Gaithersburg, Maryland 20899

Dear Dr. Ambler:

The Subcommittee on Commerce, Consumer Protection, and Competitiveness is considering H.R. 1384, legislation to require a copy-code scanner in digital audio taping devices. At the Subcommittee's hearing on May 14, there was disagreement as to the effect, if any, of the copy-code scanner system, developed by CBS, on the quality of prerecorded music.

We are requesting the National Bureau of Standards to evaluate the copy-code system. This evaluation should test whether the system would have any adverse effect on the quality of prerecorded sound or music. It is important that this evaluation be performed as expeditiously as possible since this legislation is pending before the Subcommittee. We have asked both the recording industry and members of the electronics industry opposing this legislation to cooperate with you in this test. We urge you to expeditiously consult with both sides to ensure a fair evaluation.

It is our understanding that the House and Senate Judiciary Committees have requested a similar evaluation. It is our hope that a single evaluation by the National Bureau of Standards will satisfy all requests, but we defer to your judgement on that issue.

We appreciate your attention to this request and look forward to the results. Could you please let us know approximately how long it will take to complete the test?

Thank you for your cooperation.

Sincerely,

J J Florio
James J. Florio,
Chairman
Subcommittee on Commerce, Consumer Protection, and Competitiveness

Bill Dannemeyer
William E. Dannemeyer,
Ranking Republican Member
Subcommittee on Commerce, Consumer Protection, and Competitiveness

cc: Members of the Subcommittee

[illegible]

APPENDIX B

Compact Discs Used in False Positive and False Negative Studies

Table B-1

Disc Reference Number	Disc Identification Code	Composer/Title/Artist/Company
1	CDP 7 46001 2	Pink Floyd/The Dark Side of the Moon/ Capital
2	410987	Beethoven/Symphony No. 9/ Karajan/Berlin Philharmonic/ Deutsche Grammophon
3, 31	CD 80059	Vaughan Williams/Tallis Fantasia/ Slatkin/St. Louis Symphony Orchestra/ Telarc Digital
4	MYK 37261	Debussy/La Mer; Prelude to the Afternoon of a Faun/Jeux Boulez/ New Philharmonia Orchestra/ CBS
5	410164	Prokofiev/Alexander Nevsky/ Chailly/Cleveland Orchestra/ London/Polygram Records
6	4000452	Vivaldi/Four Seasons/ Standage/Pinnock/English Concert/ Deutsche Grammophon
7	CD 80060	Beethoven/Symphony No. 5/ Ozawa/Boston Symphony Orchestra/ Telarc Digital
8	CK 40158	Judas Priest/Turbo/ Columbia
9	DMP CD 459	Manfredo Fest/Braziliana/ Digital Music Products
10	MK 7194	Wendy Carlos/Switched-On Bach/ CBS Masterworks
11	MCAD 6144	Back To The Future/ MCA Records
12	5E 507 2	The Cars/Candy-O/ Elektra

Table B-1 (con't)

Disc Reference Number	Disc Identification Code	Composer/Title/Artist/Company
13	829 179 2	The Moody Blues/The Other Side of Life/ Polygram Records
14	ARCD 8400	GTR/ Arista
15	A2 90479	Stevie Nicks/Rock A Little/ Modern Records
16	2011-2	Peter Gabriel/Security/ Geffen Records
17	CD 80094	Star Tracks/Kunzel/ Cincinnati Pops Orchestra/ Telarc Digital
18	RCD 14413	Mozart/Symphonies Nos. 40 & 41/ Chicago Symphony Orchestra/Levine/ RCA
19	412 627 2	John Williams/America, The Dream Goes On/ Philips
20	410 116 2	Dvorak/Symphony No. 9/ Chicago Symphony Orchestra/ Sir Georg Solti/ London
21	KEM 01004	Sousa/Famous Marches/ Hunsberge/Eastman Wind Ensemble/ KEM Disc
22	CD 1	CD-1 Test Disc/ CBS
23	TCD 06134	Stevie Wonder/In Square Circle/ Tamla/MCA Distributing
24	5004	Equale Brass/Bacchanales/ Nimbus Records
25	MK 42258	Bach/Concertos for Violins & Strings/ Stern, Zukerman, Schneider, Gomberg/ CBS Masterworks

Table B-1 (con't)

Disc Reference Number	Disc Identification Code	Composer/Title/Artist/Company
26	MK 37848	Rodrigo/Concierto De Aranjuez/ John Williams/Fremaux/Philharmonia Orchestra/ CBS Masterworks
27	CD 3902	The Police/Every Breath You Take/ A&M Records
28	CD 80086	Grofe/Grand Canyon Suite and Gershwin/Catfish Row; Porge & Bess/ Erich Kunzel/Cincinnati Pops Orchestra/ Telarc Digital
29	410 025 2	Gershwin/Rhapsody in Blue and Bernstein/West Side Story/ Los Angeles Philmarmonic/ Deutsche Grammophon
30	CD 80078	Copland/Appalachian Spring; Rodeo; Fanfare for the Common Man/ Louis Lane/Atlanta Symphony Orchestra/ Telarc Digital
31	Same as #3	
32	WD 1034	Liz Story/Unaccountable Effect/ Windham Hill Records
33	CDC 7 47395 2	A Bach Festival for Brass & Organ/ The Empire Brass and Douglas Major/ Angel/EMI Records
34	411 929 2	Bach, Widor, Albinoni, Franck, Mendelssohn/Toccata & Fugue Great Organ Works/London/DECCA Record
35	CD 80115	Orchestral Spectaculars/ Erich Kunzel/Cincinnati Pops Orchestra/ Telarc Digital
36	RCD 1 5303	Schubert/Arpeggione, Die Schone Mullerin, Serenade/ Galway, Moll/ RCA

Table B-1 (con't)

Disc Reference Number	Disc Identification Code	Composer/Title/Artist/Company
37	410 042 2	Chopin/Piano Concerto No. 2/ Davidovich/Marriner/ Philips/Polygram Records
38	CDC 747005 2	Bach/Violin Concertos for Violins and Strings/ Mutter/Accardo/English Consort/ Angel, EMI Records
39	CK 39936	Journey/Raised on Radio/ Columbia
40	ZK 34224	Kansas/Leftovertime/ Kirshner
41	9 24088 2	Peter Gabriel/So/ Geffen Records
42	810 002 2	Rush/Signals/ Mercury
43	ZK 34929	Kansas/Point of Know Return/ Kirshner
44	9 24098 2	Ric Ocasek/This Side of Paradise/ Geffen Records
45	9 24162 2	Aerosmith/Permanent Vacation/ Geffen Records
46	CK 40092	Barbra Streisand/The Broadway Album/ Barbra Streisand Columbia
47	400 029 2	Tchaikovsky/Symphony No. 6/ Los Angeles Philharmonic Orchestra/ Deutsche Grammophon
48	400 048 2 10	Beethoven/Violin Concerto/ Kyung Wha Chung/ DECCA/London
49	827 409 2	Pat Metheny/Watercolors/ ECM Records

Table B-1 (con't)

Disc Reference Number	Disc Identification Code	Composer/Title/Artist/Company
50	PCDI 5459	Miller/The Unforgettable/ Glenn Miller and His Orchestra/ RCA
51	MCAD 5570	Black Friday/ MCA Records
52	CAL 9926	Olivier Messiaen/Les Corps Glorieux, L'Aslension/ Calliope
53	MK 42333	Wendy Carlos/Secrets of Synthesis/ CBS Inc.
54	MK 42014	Beethoven/Symphony No. 9/ Bruno Walter/Columbia Symphony Orchestra/ CBS
55	EK 40600	Michael Jackson/Bad/ Epic
56	CDC 7 47189 2	Beethoven/Symphony No. 9/ Klemperer/Philharmonic Orchestra/ Angel, EMI Records

APPENDIX C

Contractor's Report

Appendix C: Contractor's Report -- Subjective Listening Tests

This research was supported by Purchase Order 43NANB800479 between the National Bureau of Standards (NBS) and the University of Michigan, Principal Investigator, Irwin Pollack. The contract monitor for NBS was Daniel R. Flynn.

The opinions expressed in this Appendix are those of the contractor and do not necessarily represent those of the National Bureau of Standards.

Summary

Subjective listening tests were carried out to examine the extent to which listeners could discriminate between unencoded music and music that had been encoded using the CBS Copy-Prevention Encoder. Listening materials consisted of excerpts of music from commercially-available stereo compact discs plus, for the serial tests described below, a few samples produced by a music synthesizer. The music was either passed through the Encoder, or by-passed the Encoder.

For the "serial tests," listeners attempted to identify whether repeated pairs of sequentially-presented musical passages were selected from the same source (either encoded or unencoded), or from different sources (one encoded, one unencoded). Large differences in average discrimination performance were associated with different audio materials in the serial listening tests. Based upon the combination of earphone and loudspeaker testing, 5 of 20 compact disc selections and 2 of 4 synthesizer examples were discriminated at a level significantly higher than chance. The results suggest that the expected level of discrimination is highly sensitive to the specific materials examined. Differences among listeners tended to be smaller than differences among materials.

In the "parallel tests," listeners could switch at will between parallel pairs of tracks of encoded and unencoded selections in an attempt to determine which material differed from the unencoded material on a pair of reference tracks. Substantially better performance was achieved in the parallel tests than in the serial tests. Four of 15 listeners achieved discrimination levels significantly higher than expected by chance (one listener achieving a perfect score of 10 of 10 selections); and 4 of 10 selections yielded discrimination levels significantly higher than expected by chance.

Introduction

The Congress of the United States asked the National Bureau of Standards (NBS) to examine several aspects of the copy-prevention system developed by CBS Records. Among the questions to be addressed by NBS was whether the quality of recorded music is diminished by the use of the copy-prevention system. The portion of this system that potentially could affect musical quality is the Encoder; the Decoder is not involved in this aspect of the NBS evaluation of the Copy-Prevention System.

In preliminary discussions, the Principal Investigator and staff members of the NBS Acoustic Measurements Group and the NBS Statistical Engineering Division discussed the interpretation of the Congressional directive. An obvious test was whether or not listeners could tell when the Encoder was operating. We call this a "chocolate-versus-vanilla" test. A possible approach would have been to present musical selections with and without the Encoder and to ask listeners to identify which one was produced by the Encoder. Several NBS staff members argued strongly against that approach. Among their reasons was that it would not be possible to give sufficient training to the listeners to enable them to develop a proper 'taste' for chocolate and vanilla, i.e., to appreciate encoded versus unencoded materials.

An alternative approach to the question asked by Congress was to take the position that aural discrimination of any (non-trivial) recognizable departure, due to the encoding process, from an original recording represents a degradation of the quality of that recording. (An example of a trivial departure from the original performance is a change in overall level which can easily be compensated.) This interpretation was taken, so that in this study, instead of asking listeners to identify the presence of the Encoder, they were asked whether musical selections were reproduced from the same source or from different sources, without being asked to identify the nature of the source. This decision is crucial to the tests which follow.

Assumptions

1. A broadly-based scientific examination of the Encoder would involve a detailed examination of the effects of selected variables on the recognition of the Encoder.

2. The discrimination of a non-trivial modification of a recorded performance constitutes a perceived degradation in quality.

3. In the tradeoff between the length of the testing materials and the variety of materials, it was decided to examine a wide variety of recorded materials and relatively short musical samples.

4. In the tradeoff between professional experience and "young" ears (which would be less likely to suffer from hearing loss), the decision was made not to limit the study to young subjects, but to include experienced listeners, especially those with a background in music and in sound recording.

5. With regard to the question of whether (1) to utilize a random selection of music and subjects that would, in some statistical sense, be "representative" of larger populations or (2) to make "critical" selections of music (where the effects of the Encoder might be more audible) and of listeners (who would be more likely to hear such effects), the decision was made to pick critical selections and subjects.

6. It was decided not to give the subjects specific listening training as to the effects of the Encoder, prior to testing.

7. Although discrimination of the effects of the Encoder might have been more likely if the subjects were provided with visual displays of the results of physical measurements, it was determined that all tests would be confined to auditory listening.

8. To ensure impartiality of testing administration, all tests would be run 'double-blind' -- without the test administrator having knowledge of the key.

9. A difficult decision was made with respect to the loudness listening level for the serial tests, for which the subjects were presented with a constant listening level rather than being permitted individual adjustment to a preferred listening level.

Discussion of Assumptions, Objections, Tradeoffs

The following items correspond to the numbers of the assumptions given above:

1. Under the time and financial constraints of the Project, a broadly-based basic scientific examination of the Encoder was deemed impossible.

2. Assumption 2 was crucial to the research. Insufficient time was available to train listeners to hear any effects that the Encoder might have on a wide variety of types of music. In other words, a chocolate-versus-vanilla test would be impossible unless listeners could be provided with a fuller understanding of the 'tastes' of chocolate versus vanilla.

3. Since certain classes of music reproductions might be more sensitive than others to the action of the Encoder, the decision was made to utilize a range of musical selections so as to ensure adequate coverage.

   Either of two extreme cases might represent an appropriate listening test. One could present isolated individual notes in which Encoder action was identified -- this would completely deprive the listener of an adequate musical setting or context. One could present long musical passages in which an isolated Encoder action occurred -- this would place an intolerable burden on the listener's auditory memory, asking him in effect, to "locate a needle in a giant haystack." The compromise that was made for the serial study was to ask the listener to "locate a needle in a small haystack" by presenting selections that were nominally 5 seconds in length.

4. The original plan included using music performance and music recording majors from local universities as listeners. The tests were delayed so that the projected starting time fell in the middle of the academic Christmas vacation. Meanwhile, the response from the audio engineering

community was so enthusiastic that the decision was made not to return to the university students. It was obviously important to take advantage of the critical listening skills of experienced audio engineers and musicians. Even more importantly, many of these experienced professionals have devoted their professional careers toward the improvement of music reproduction. It would not be possible to duplicate the motivation and intrinsic interest of this group by means of financial inducements upon a less concerned population.

Because of reports that audio engineers may suffer hearing loss in the active frequency region of the Encoder (4 kHz), an audiometric examination was carried out on each of the subjects in the serial listening tests.

5. The NBS statistician on the study pointed out that the project lacked the time and financial resources to achieve a "representative" sampling of recorded music and of listeners. Since it made little sense to test the performance of the Encoder when the Encoder was not engaged or to test the Encoder on unmotivated or unskilled listeners, it was decided to opt in the direction of a "critical" sampling of music and of listeners.

6. Specific training was not given for several reasons. Within the time available for testing volunteer listeners, additional training time would have reduced the range of examples for testing. There was concern that to give extensive training as to the effects of the Encoder might be construed as biasing the study. And, finally, there was reluctance to serve up, as practice examples, the most easily detected selections.

7. In examining compact discs for Encoder action, NBS staff utilized physical measurements, computer analyses, and visual displays to screen musical candidates for later aural study and to identify specific effects of the Encoder on particular musical passages. It was found that the Encoder was often most active under conditions where aural identification was very difficult, e.g., when an entire orchestra was producing music. Since music listeners to not normally have spectral analysers available, since it would be difficult do teach listeners to understand the implications of spectral displays, and since to make sophisticated instrumentation available to the listeners might be construed as biasing the study in the direction of easy detection of any audible effects that the Encoder might have, it was decided that the listeners would have to make do with their ears.

8. Because of the sensitive nature of the tests, test impartiality was considered to be a primary requirement. The serial tests were run by the Principal Investigator, who did not know the answers. After the parallel-test subjects were trained in the use of the multi-track tape recorder, they were left alone to do their listening.

9. A single listening level was selected for the serial tests to ensure uniform testing and to avoid unacceptably loud levels.

Limitations of the Present Study

The same-different sequential procedure is not without difficulty. A heavy burden is placed on auditory memory. A listener may detect a large difference with high confidence. However, when no difference is detected, the listener is unsure whether the items are actually the same or whether there exists a difference which is indistinguishable.

The testing program did not allow for extended specialized training. There is no doubt that, with sufficient training, practice, and motivation, incredible individual feats may be attained. Indeed, recognized critics are often respected for the development of highly specific talents. For the purposes of this study, however, it is surmised that moderate amounts of additional training probably would have produced only modest improvements in performance.

Experimental Approach: Serial Tests

Listening Materials

The selection of the test materials for the listening tests is described in section 4.4. Specific compact discs were recommended by the Home Recording Rights Coalition and the Recording Industry Association of America. With the aid of spectrum analyzers, computer monitoring, visual displays, and a read-out of the instanteous behavior of the Encoder, significant periods of Encoder activity could be located. Aided by this information, NBS staff then selected examples where significant aural changes appeared to take place. In this way, the sophisticated instrumentation served to screen potential candidates for testing, but the final selection was based upon aural identification.

Listeners

Listeners were recruited among members of the Washington-Baltimore chapter of the Audio Engineering Society, among local radio stations and recording studios, and among National Bureau of Standards employees who were knowledgable about music.

Recording Procedure

The recording procedures for the test materials are detailed in Appendix F. In essence, the test materials were derived as follows: the first generation consisted of a 5-10 second sample of music with manual fade-in and fade-out; the digital sample was converted to an analog signal; that signal was both passed through the Encoder and by-passed the Encoder; and a second generation digital sample was obtained by reconversion of the analog signals. This procedure ensured that both encoded and unencoded samples were treated identically except for passage through the Encoder.

Audio Equipment and Calibration

The earphones and loudspeakers used for this study, along with their calibration, are detailed in Appendix G. In essence, high quality professional equipment was employed throughout the tests. Of special concern

was the selection of loudspeakers. NBS staff selected a two-way speaker system with a crossover frequency below 2000 Hz, rather than a three-way speaker system with a crossover frequency in the active region of the Encoder. In this way, the phase distortion of the loudspeaker crossover network did not interact with that of the Encoder.

Test Procedure

Listeners heard sequences of six presentations of the same musical sample. A trial consisted of a given A-B pair repeated three times, e.g., A_1B_1, A_2B_2, A_3B_3. All of the A samples and all of the B samples were the same within a given trial. Across different trials, the A or B samples were equally-often encoded (E) or unencoded (here called direct (D)). All four possibilities were available: EE, DD, ED, and DE. The correct response was 'same' to the EE and DD pairings and 'different' to the ED and DE pairings. Note that the listener did not have to identify whether a sample was or was not encoded.

The intra-pair A_1B_1 duration was one second; the inter-pair B_1A_2 duration was three seconds. Each test sequence was preceded by the announcement "Here is test pair x;" and was terminated by the announcement "Please encircle your response in slot x." For each musical selection, the entire sequence of announcements, music, and pauses averaged about 56 seconds in duration. Four practice examples were employed to introduce the test procedures. Discussion was held after the four examples to ensure understanding of the test procedure.

Experimental Results: Serial Tests

Qualitative Observations

On the basis of articles in leading semi-technical audio magazines, many of the listeners anticipated that gross and blatant changes would be introduced by the Encoder. As dedicated professionals in the recording and musical arts, they were obviously concerned. They were eager to hear music through the Copy-Code system, freely volunteering time from their busy professional careers in order to participate in the tests. Hardly any degree of financial reward with disinterested observers could have yielded the equivalent motivation. This point is significant because negative findings can easily be obtained from the best-intentioned experimental design, if the listeners are poorly motivated. Regardless of the strong attitudes of the listeners, the double-blind procedure and experimental design served to safeguard the study.

At the conclusion of the listening session, many of the listeners remarked that the effect of the Encoder was substantially more subtle than they had expected.

Qualitative observations such as those cited above, of course, are not adequate substitutes for more controlled, quantitative test scores. Nevertheless, they suggest that the effect of the Encoder can be subtle and that quantitative tests might not demonstrate high discrimination scores.

Quantitative Observations: Musical Selections

Substantial differences in discrimination performance were associated with different musical selections. Table C-1 presents the average scores over 61 listeners employing earphones for the entire set of 24 selections (20 music and four synthesizer examples). The range of correct identification scores is 41.4% to 91.8%. Marked with plus (+) signs are four of 24 examples yielding an average performance which was statistically significantly better than an expected chance performance of 50%. (Statistical significance is based upon the binomial probability distribution at the 95% confidence level; see section 4.5). Marked with a minus sign (-) is one of twenty-four examples that yielded an average performance significantly poorer than chance. Similarly, table C-2 presents the average scores over twenty-three listeners employing loudspeakers for the entire set of twenty-four selections. The range of correct identification scores is 40.2% to 87.0% with 2 of twenty-four selections yielding scores significantly greater than expected by chance. Figure C-1 presents the cumulative distribution of average discrimination scores over the twenty-four selections for earphone (E) and loudspeaker (S) listeners.

In the analysis in section 4.5, earphone and loudspeaker listeners (including 3 listeners who used earphones part of the time and loudspeakers part of the time) were combined into a single group. As a result of the greater statistical sensitivity of the combined group, eight selections (six compact disc and 2 synthesizer) exceeded the expected chance level by a statistically significant margin, and one selection was significantly lower than expected by chance. The correlation over musical selections between scores for earphones listening and for loudspeaker listening was moderate (r = 0.80), as shown in figure C-2. The correlation is heavily weighted by one extreme case.

The synthesized B_7-F_7 chord yielded the highest discrimination score. Application of a statistical test for extreme cases (the Dixon test) shows that specific example to be an atypical, or extreme, case from other examples. This result does not mean that we can reject that example, but, rather, that the example appears not to have been drawn from the same normal distribution as the other scores. In partial summary, test discrimination scores for the Encoder are strongly dependent upon the specific selection tested.

On average, the same-different test scores across selections yielded discrimination significantly better than chance. Specifically, of 61 earphone listeners, 45 scored better than 50 percent correct, a result expected by chance about 1 in 10,000 times.

Among the 24 selections, the median proportion for correct responses was 0.522 for loudspeaker listening and 0.533 for earphone listening. Whereas the median scores are only marginally higher than chance, the scores for individual selections may demonstrate substantial discrimination.

TABLE C-1

SCORE SUMMARY BY SELECTION FOR EARPHONE LISTENERS

Selection Name	Number of Responses	Percent Correct
Debussy	244	54.1
Barber	244	50.4
Beethoven chord	244	54.5
Beethoven chorus (-)	244	41.4
Bizet	244	55.7
Wilson	244	52.9
Streisand	244	50.8
Beethoven Triangle	244	54.9
McGovern	244	53.3
Weather	244	43.9
Copland	244	54.5
Bernstein	244	50.4
Pink Floyd	244	51.6
S1 (+)	244	91.8
Prokofiev (+)	244	73.4
S2	244	54.1
Metheny	244	52.5
S3	244	53.3
S4	244	49.2
Messiaen	244	54.9
Bach	244	52.5
Respighi (+)	244	61.1
Handel	244	49.2
Ravel (+)	244	57.0

TABLE C-2

SCORE SUMMARY BY SELECTION FOR LOUDSPEAKER LISTENERS

Selection Name	Number of Responses	Percent Correct
Debussy	92	49.0
Barber	92	49.0
Beethoven chord	92	54.3
Beethoven chorus	92	47.8
Bizet	92	52.2
Wilson	92	43.5
Streisand	92	41.3
Beethoven Triangle	92	50.0
McGovern	92	51.1
Weather	92	52.2
Copland	92	54.3
Bernstein	92	44.6
Pink Floyd	92	40.2
S1 (+)	92	87.0
Prokofiev	92	60.0
S2	92	55.4
Metheny	92	56.5
S3	92	60.0
S4	92	52.2
Messiaen	92	60.0
Bach	92	50.0
Respighi (+)	92	68.5
Handel	92	51.1
Ravel	92	52.2

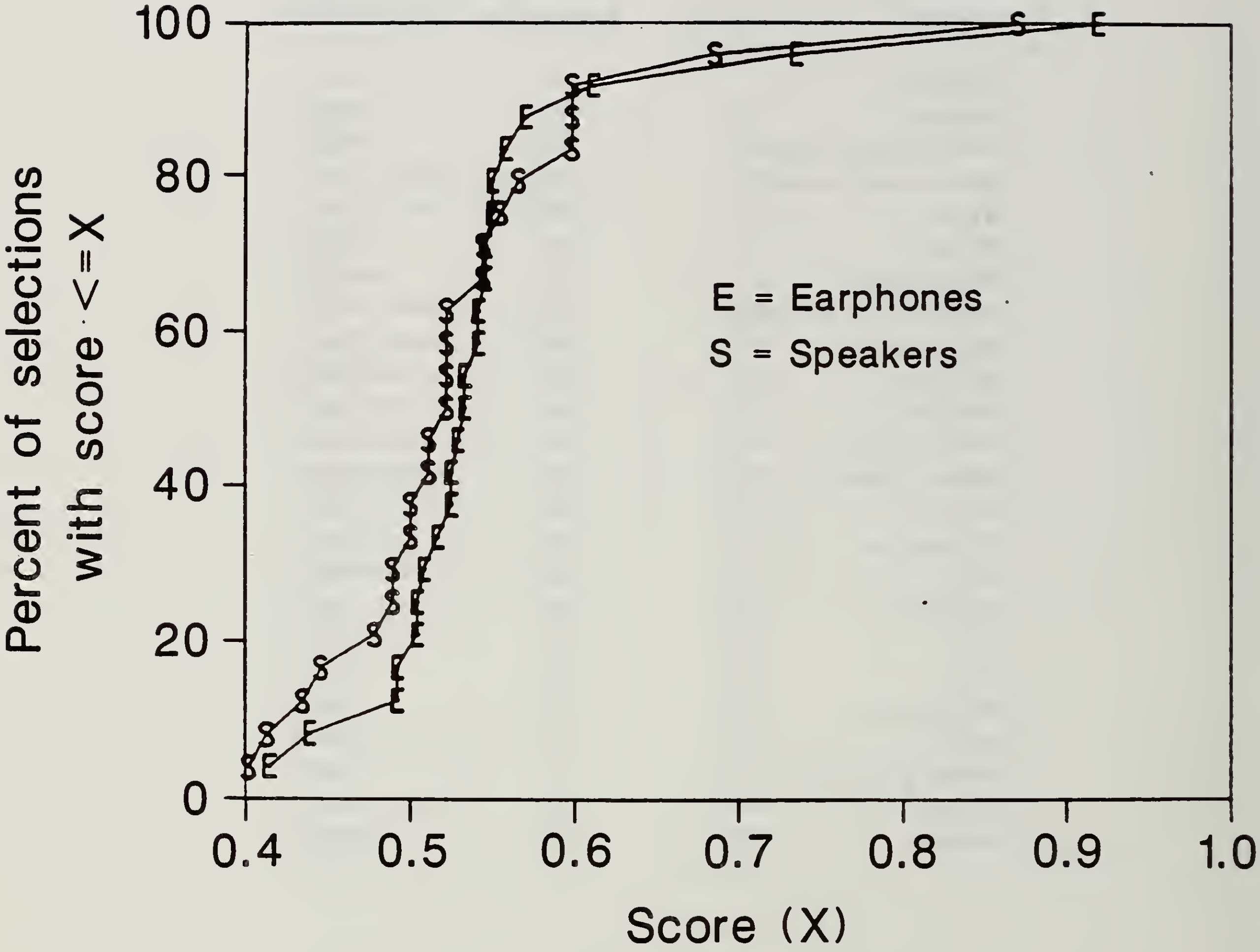


Figure C-1 Distribution of same-different test scores over 24 selections, including four synthesizer selections. Results for earphone and loudspeaker listening shown separately.

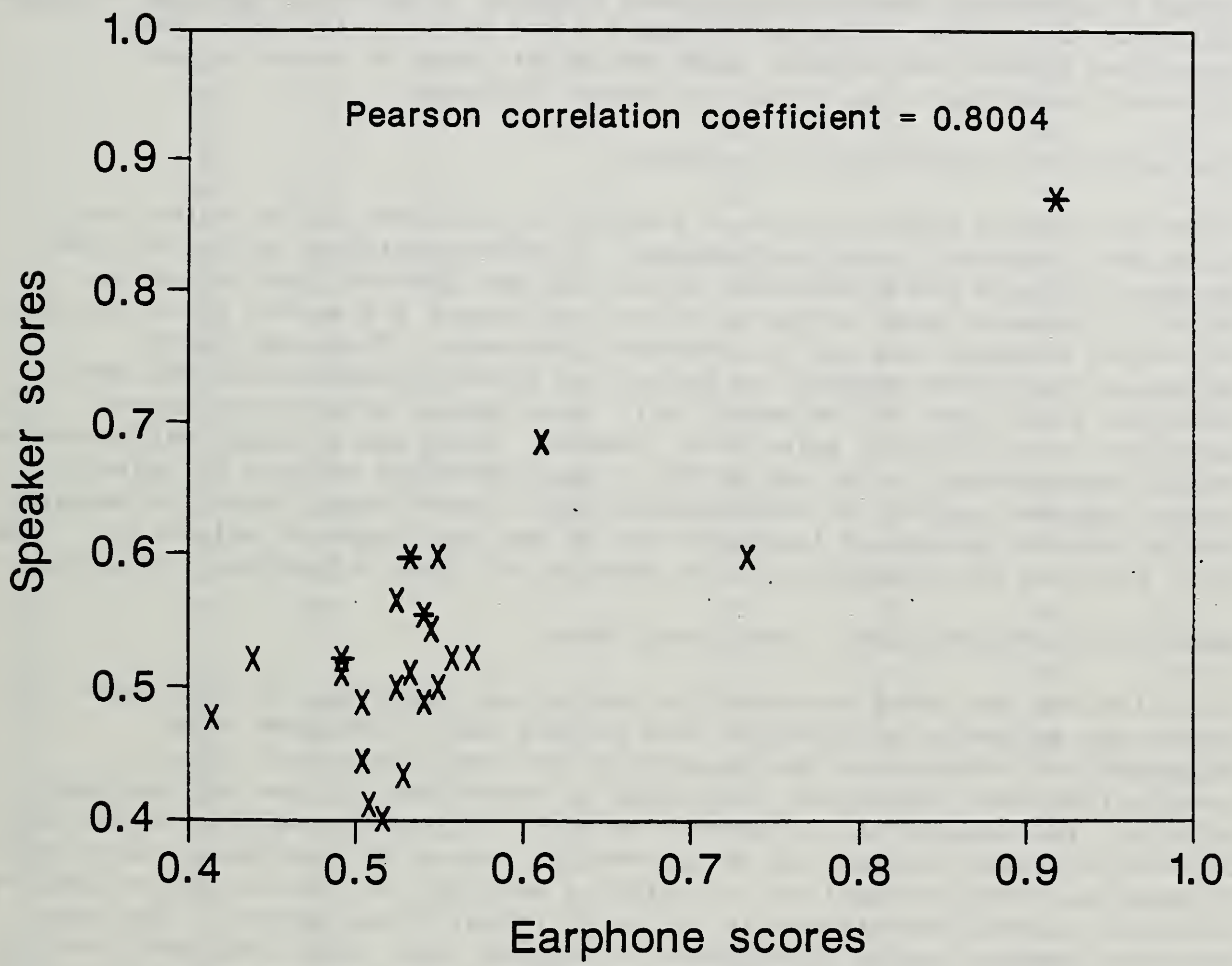


Figure C-2 Comparison between same-different test scores over 24 selections, including four synthesizer selections plotted as stars. Results for earphone and loudspeaker listening are plotted on separate axes.

Quantitative Observations: Listeners

As indicated in the last column of tables C-3 and C-4, 10 of the 61 earphone listeners, and 3 of the 23 loudspeaker listeners, scored at a discrimination level significantly higher than expected by chance in the serial test. The highest discrimination score achieved was 63.5% for loudspeaker listening.

Figure C-3 presents the cumulative rank ordering of scores associated with the individual listeners. Figures C-1 and C-3 are drawn to the same scale. Comparison between the figures shows the wider range of scores across different selections than across different listeners.

Interaction of Selections and Listeners.

There was concern that statistical pooling of listeners and/or selections would mask inherently good performance. A median empirical split was made between listeners and selections, where each was divided into two groups. Table C-5 presents mean scores with the resulting 2 x 2 median split for the 20 musical examples and for the earphone listeners. There was little difference in scores between the better and poorer listeners for the more difficult selections (47.9% and 51.6%). Both groups scored near chance. Among the less difficult selections, however, there was a larger difference in average performance (52.6% and 61.9%). The individual entries of table C-5 can be compared against an independence model, based simply upon the marginal sums of the two groups of listeners and of the two groups of selections. That model predicts the separate tabular enteries of table C-5 within 1.1 percent.

Quantitative Observations: Additional Notes

Each listener was asked to supply his or her age and years of experience, to report any incidence of tinnitus, and to have their audiogram taken. (Biographical information was incomplete for some listeners.) Also, correlations were run across test items to search for fatigue and learning effects. The results of a step-wise multiple regression analysis showed near-zero correlations between scores against age, years of experience, incidence of tinnitus, and hearing loss. (A similar analysis of factors against hearing loss will be reported separately by the Principal Investigator.) The near-zero correlations can be interpreted to suggest that, under the conditions of testing employed, we would not expect large differences among other groups of listeners. In addition, near-zero correlation with successive test items suggests that learning and/or fatigue effects either balanced out or were not strong.

Experimental Approach: Parallel Tests

Introduction

In the same-different serial tests of the previous sections, the listeners heard nominally 5-second musical samples. It might be argued that a skilled listener would better be able to identify the effects of the Encoder by exploring longer musical selections for critical instances of his or her own choosing. To a recording engineer, the preferred procedure is a 'hands-on'

TABLE C-3

SCORE SUMMARY BY SUBJECT FOR EARPHONE LISTENERS
ALL 24 SELECTIONS (96 ITEMS) USED

Row Index	Station	Age	Experi- ence	Hearing Loss	Tinnitus	Subject	Number of Responses	Percent Correct*
1	E1	27	8	10	Y	1	96	60.4(+)
2	E2	42	15	15	N	2	96	57.3
3	E3	59	35	40	N	3	96	55.2
4	E4	32	12	3	Y	4	96	50.0
5	E5	39	29	3	N	5	96	58.3
6	E1	69	38	48	N	6	96	54.2
7	E2	30	10	8	N	7	96	53.1
8	E3	41	10	30	Y	8	96	59.4(+)
9	E4	33	8	3	N	9	96	55.2
10	E5	38	15	25	N	10	96	55.2
11	E1	31	20	18	N	11	96	52.1
12	E2	25	5	18	N	12	96	53.1
13	E3	52	30	10	Y	13	96	57.3
14	E4	43	30	25	N	14	96	51.0
15	E5	38	6	13	N	15	96	59.4
16	E1	37	5	5	N	16	96	51.0
17	E2	38	6	13	Y	17	96	55.2
18	E3	28	.	5	N	18	96	56.3
19	E1	27	7	5	N	19	96	47.9
20	E2	51	.	8	Y	20	96	59.4
21	E3	34	12	5	N	21	96	58.3
22	E1	26	7	28	Y	22	96	52.1
23	E2	23	4	10	N	23	96	61.5(+)
24	E3	24	4	10	N	24	96	58.3(+)
25	E4	23	3	13	Y	25	96	50.0
26	E1	35	13	8	N	26	96	49.0
27	E2	50	30	27	Y	27	96	61.5(+)
28	E3	22	.	5	N	28	96	51.0
29	E4	26	3	5	N	29	96	53.1
30	E1	42	12	10	N	30	96	56.3
31	E2	39	20	15	N	31	96	52.1
32	E3	41	25	45	N	32	96	51.0
33	E4	26	5	10	Y	33	96	61.5(+)
34	E5	24	4	13	N	34	96	51.0
35	E1	36	5	5	N	35	96	53.1
36	E2	47	15	.	Y	36	96	50.0
37	E3	35	15	8	N	37	96	56.2
38	E4	36	18	3	N	38	96	54.2
39	E5	40	.	5	N	39	96	55.2
40	E1	48	10	8	N	40	96	58.3(+)
41	E2	37	15	15	Y	41	96	53.1
42	E3	23	3	3	N	42	96	56.2
43	E4	25	6	15	N	43	96	54.2

[TABLE C-3, cont'd]

44	E5	49	20	10	N	44	96	52.1
45	E1	36	4	8	N	45	96	46.9
46	E2	45	6	15	N	46	96	59.4(+)
47	E3	65	25	23	N	47	96	51.0
48	E2	62	35	13	N	50	96	50.0
49	E3	46	27	15	N	51	96	58.3
50	E3	26	7	25	Y	55	96	55.2
51	E1	36	14	8	Y	57	96	63.5(+)
52	E2	34	7	15	Y	58	96	55.2
53	E3	43	7	30	N	59	96	54.2
54	E1	48	11	13	N	60	96	58.3
55	E2	40	1	13	N	61	96	53.1
56	E3	24	1	3	N	62	96	49.0
57	E4	25	3	3	N	63	96	56.3
58	E1	38	11	5	N	64	96	60.4(+)
59	E2	29	16	5	Y	65	96	53.1
60	E3	26	5	5	Y	66	96	56.3
61	E4	24	6	5	N	67	96	57.3

*Scores significantly higher than chance ($p < 0.025$) are marked "(+)."

TABLE C-4

SCORE SUMMARY BY SUBJECT FOR LOUDSPEAKER LISTENERS
ALL 24 SELECTIONS (96 ITEMS) USED

Row Index	Station	Age	Experi- ence	Hearing Loss	Tinnitus	Subject	Number of Responses	Percent Correct
1	S1	36	15	10	N	80	96	50.0
2	S2	24	4	3	N	81	96	52.1
3	S1	39	22	30	N	82	96	58.3
4	S2	65	45	38	N	83	96	57.3
5	S1	53	30	10	N	84	96	46.9
6	S2	28	14	.	N	85	96	61.5(+)
7	S1	66	45	45	Y	86	96	61.5(+)
8	S2	.	.	.	.	87	96	55.2
9	S1	27	2	5	N	88	96	51.0
10	S2	32	10	60	Y	89	96	52.1
11	S1	38	15	20	N	90	96	55.2
12	S2	33	12	5	N	91	96	52.1
13	S1	44	20	3	N	92	96	57.3
14	S2	55	12	33	Y	93	96	46.9
15	S1	25	12	3	Y	94	96	54.2
16	S2	29	12	10	N	95	96	52.1
17	S1	33	12	15	N	96	96	52.1
18	S2	23	7	5	N	97	96	50.0
19	S1	49	34	8	N	99	96	49.0
20	S2	28	7	8	N	100	96	45.8
21	S2	37	15	13	Y	101	96	61.5(+)
22	S1	55	25	10	Y	103	96	56.3
23	S2	74	40	53	Y	104	96	50.0

*Scores significantly higher than chance ($p < 0.025$) are marked "(+)."

TABLE C-5

MEDIAN SPLIT ANALYSIS OF AVERAGE SCORES
ACROSS EARPHONE LISTENERS (n=61) AND MUSICAL SELECTIONS (n=20).

Selections

		≤Median (n=10)	>Median (n=10)
Listeners	≤Median	47.9%	52.6%
	>Median	51.6%	61.9%

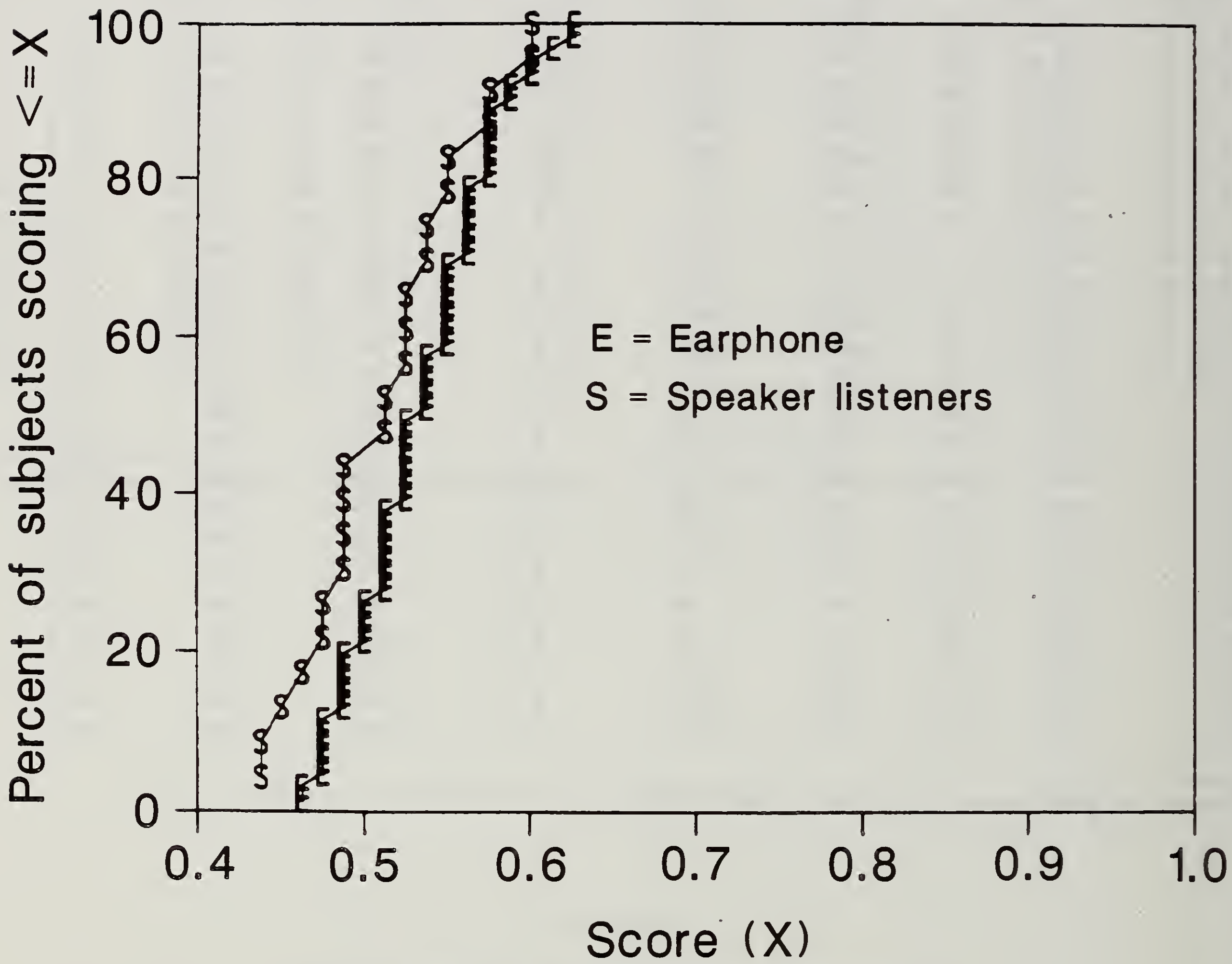


Figure C-3 Distribution of same-different test scores over listeners for the 24 selections. Results for earphone and loudspeaker listening shown separately.

approach, providing unrestricted access to the music. The tests of this section, called parallel tests, provided much less restricted access than in the case of the serial tests. These parallel tests also exploited differences in quality between the two stereo channels. The Encoder's action is based upon the average of signals in the two stereo channels. In some instances, because of the differences between the two stereo channels, the Encoder will operate more effectively upon one of the two stereo channels than the other. While such cues were also available in the serial tests, they were probably less effective because of the shorter duration of the signals.

Recording Procedure

A multi-track digital tape recorder was used, with ten selections of 29 to 75 second duration with three parallel pairs of stereo tracks: an unencoded stereo pair of reference channel tracks and, in a scrambled order, one pair of unencoded test tracks and one encoded pair of test tracks. The unencoded reference track-pair was identified in advance to the observer. The observer did not know which of the pairs of test track-pairs was encoded. The selections encompassed the shorter 5-10 second selections of the serial tests. The musical selections are listed in Appendix E.

Test Procedure

The task of the listener was to select which one of the two pairs of test tracks contained encoded material, i.e., which differed from the reference unencoded track-pair. The listener could switch at will between the reference track-pair and the two test track-pairs. Earphone listening was employed for all tests. Only one track-pair was audible at any instant. Time tabs were available for repeated presentation of portions which appeared to be especially relevant to the listener. The listener could adjust the overall signal level to a preferred level.

The listeners were generally recording engineers, audio equipment designers, or professional musicians.

Qualitative Observations

The listeners were encouraged to transcribe the basis for each of their answers into a tape recorder. Some of their remarks were detailed, pointing to the exact tape location of specific events. It is interesting to speculate whether the cumulative experience of providing the remarks of earlier expert listeners to later listeners would have provided an accurate consensus.

Some of the remarks reflected extraordinary insight into the operation of the Encoder. For example, "It sounded to me that something of lower frequency was causing the filter to go in and out on an amplitude-dependent basis." Or "I had assumed I would hear a simple timbral shift, if anything, and that the timbral shift would be comparatively minor. Instead, in some instances, it seems to me that the actual impact is quite pronounced and because it's variable, it's extremely troubling." Some of the remarks also reflected differences in stereo location.

Because some of the listeners scored significantly poorer than expected by chance, the listeners were interviewed by telephone (at a later date) about their understanding of the instructions. All listeners correctly responded that their task was to identify the encoded selection, or the selection that differed from the unencoded reference.

Quantitative Observations

The number of correct selections, out of 10, is presented for each listener and for each selection in table C-6. By chance, a score of 10 correct of 10 selections would be expected 1 out of a 1000 times, a score of 9 or more correct would be expected 1 out of a hundred, and a score of 8 or more out of 10 would be expected somewhat more than 5 percent of the time ($p = 0.055$). Employing this latter criterion, one can say that 4 of the 15 listeners scored significantly higher, and 1 of the listeners scored significantly lower, than expected by chance. The average score, 62.3 percent, is substantially higher than that achieved with the main body of serial listening tests.

Similarly, by chance, a score of 12 or more correct of 15 listeners would be expected by chance about 2 times in a 100, and 11 or more correct of 15 would be expected about 6 times in a 100 ($p = 0.059$). Employing this critereon, we can say that 4 of the 10 selections scored significantly higher than chance response. No selection was perfectly discriminated.

Interpretation of Results

Within a wide body of results, protagonists can pick and choose selected features which support their previously-announced positions. Protagonists can also advocate test procedures which might support their positions.

Before commenting upon the results, let us consider what constitutes a behaviorally relevant test. For example, equipping each observer with a spectrum analyzer and visual display might have resulted in perfect discrimination. Most neutral observers, however, might be expected to judge such tests as behaviorally irrelevant to the aural detection of the effects of the Encoder.

Along the same lines, a narrow bandpass filter could have screened away all frequencies outside of the active region of the Encoder. With such a technique, the action of the Encoder might have been easily detected by aural means alone. Most neutral observers, however, might be expected to judge such test to be irrelevant to the aural detection of the Encoder for wide-band music in a normal listening situation.

A test also might have been devised to take special advantage of the human binaural system. To one ear would be fed the left unencoded channel of a musical selection. To the right ear would be fed either the right unencoded channel of that selection or the right encoded channel. Alternatively, to the right ear, we would feed either the left encoded channel of that selection or the left unencoded channel. Had the binaural test display yielded substantially higher scores, would this test have qualified as a reasonable test of the Encoder? The neutral observer might have somewhat more trouble

Table C-6 Identification of encoded channel in the parallel-test format

Selection	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	Fraction Correct
	Subject															
1. Bizet	C	•	C	C	C	w	C	C	w	w	w	w	C	C	w	8/14
2. Streisand	w	C	C	C	w	C	C	w	C	C	C	C	w	C	C	11/15
3. Beethoven Triangle	C	•	C	C	C	w	C	w	C	C	w	C	w	C	C	10/14
4. Copland	w	w	C	C	C	w	C	w	C	w	w	C	w	w	C	7/15
5. Metheny	C	C	C	C	C	C	C	C	C	C	w	w	C	w	C	12/15
6. Prokofiev	C	w	w	C	C	C	C	C	C	C	w	C	C	C	C	12/15
7. Bernstein	C	w	C	C	w	C	C	w	w	C	C	C	C	C	C	11/15
8. Messiaen	C	•	C	C	w	C	w	w	C	w	w	w	w	C	C	7/14
9. Bach	w	•	C	C	C	w	C	w	C	w	w	C	C	w	C	8/14
10. Respighi	w	w	w	C	C	w	C	w	w	w	w	w	w	C	C	5/15
Total Correct	6	2	8	10	7	5	9	3	7	5	2	6	5	7	9	

C = Correct

w = Wrong

• = No Response

Average percent correct: 100 × (91/146) = 62.3

with rejecting this wide-band aural test procedure than the aforementioned test procedures.

Still along the same lines, the parallel 'hands-on' test provided more degrees of listening freedom than the serial same-different tests. Despite the gross differences between the two test procedures, the average test scores were not radically different from each other.

Even within the recognized limitations of the present tests, one would hope that neutral observers would judge the present tests as behaviorally relevant to the aural detection of the Encoder.

Where are we? I reject the extreme position that the Encoder's action is so evident and pervasive that it will be immediately recognized by unsophisticated listeners. However one feels about the principle of the Encoder, one must concede that its effect is extremely subtle. I also reject the extreme position that the Encoder's action is so benevolent that it cannot be detected. Significant discrimination was achieved in the serial tests with several selections. Even the scores of the best discriminated example, however, deviated substantially from perfect detection. The latter point is important because one claim made for employing the Encoder is that it can be by-passed in the recording studio when it "obviously" intrudes upon the music.

The results may be interpreted to follow the intermediate position of a "good news -- bad news" scenario. On the one hand, the average level of discrimination, especially in the serial test procedure, was only marginally higher than chance. On the other hand, averages can be misleading. Individual listeners and individual selections, especially in the parallel test format, yielded substantial discrimination of the Encoder's activity.

Because of individual differences among listeners, among musical selections, and among test procedures, and because improved encoding procedures might ultimately be developed, I would strongly urge that original master recordings be preserved in a unencoded mode.

Credits

The listeners in these tests must be singled out for special attention. They were unpaid volunteers who took time from their busy careers to participate in the tests. Their enthusiastic participation is exemplified by their complete turnout on a Saturday in the midst of a Washington snow storm which had closed down all government offices for a three-day weekend. It was a singular and rewarding honor and privilege to work with such dedicated professionals.

Attachments

The attachments on the next two pages are the subject answer sheets for the serial and parallel listening studies, respectively.

NBS COPY-CODE LISTENING PROJECT
Sequential Test

Name: ______________________________ Date: __________________

Listening Station: ___________________ Time: __________________

S = Same D = Different

1.	S	D	25.	S	D	49.	S	D	73.	S	D
2.	S	D	26.	S	D	50.	S	D	74.	S	D
3.	S	D	27.	S	D	51.	S	D	75.	S	D
4.	S	D	28.	S	D	52.	S	D	76.	S	D
5.	S	D	29.	S	D	53.	S	D	77.	S	D
6.	S	D	30.	S	D	54.	S	D	78.	S	D
7.	S	D	31.	S	D	55.	S	D	79.	S	D
8.	S	D	32.	S	D	56.	S	D	80.	S	D
9.	S	D	33.	S	D	57.	S	D	81.	S	D
10.	S	D	34.	S	D	58.	S	D	82.	S	D
11.	S	D	35.	S	D	59.	S	D	83.	S	D
12.	S	D	36.	S	D	60.	S	D	84.	S	D
13.	S	D	37.	S	D	61.	S	D	85.	S	D
14.	S	D	38.	S	D	62.	S	D	86.	S	D
15.	S	D	39.	S	D	63.	S	D	87.	S	D
16.	S	D	40.	S	D	64.	S	D	88.	S	D
17.	S	D	41.	S	D	65.	S	D	89.	S	D
18.	S	D	42.	S	D	66.	S	D	90.	S	D
19.	S	D	43.	S	D	67.	S	D	91.	S	D
20.	S	D	44.	S	D	68.	S	D	92.	S	D
21.	S	D	45.	S	D	69.	S	D	93.	S	D
22.	S	D	46.	S	D	70.	S	D	94.	S	D
23.	S	D	47.	S	D	71.	S	D	95.	S	D
24.	S	D	48.	S	D	72.	S	D	96.	S	D

NBS COPY-CODE LISTENING PROJECT

PARALLEL TESTS

Name: ______________________________ Date: __________________

Listening Station: ____________________ Time: __________________

ENCODED MATERIAL

1. A B

2. A B

3. A B

4. A B

5. A B

6. A B

7. A B

8. A B

9. A B

10. A B

APPENDIX D

Sponsor's Lists of Recorded Material

APPENDIX D: SPONSORS' LISTS OF RECORDED MATERIAL

The National Bureau of Standards invited the two sponsors to submit suggestions for recorded material to be included in the listening tests. These tests are reproduced in this appendix.

Home Recording Rights Coalition

The first 32 selections of those below were sent to NBS on October 6, 1987. Many of these selections were reviewed as to their suitability for inclusion in the listening tests. The remaining selections in the HRRC list were sent to NBS on December 10 and 11, 1987. Although these latter selections were received too late for consideration for the listening studies, a listing of them is enclosed here for completeness.

The following text is quoted from the October 6 transmittal letter from HRRC to NBS:

> "Attached is a list of selections from 32 compact discs which, based on screening done for HRRC, we have reason to believe illustrate the distortive effects of the CBS "copy-code" system.
>
> "Our list of suggestions is not, by any stretch of the imagination, exhaustive. It was compiled by David Ranada, Technical Editor of High Fidelity Magazine, after auditions performed through a filter believed to resemble the filter used by CBS in a Congressional hearing on April 2, 1987. The characteristics of the encoder CBS submitted to the NBS may differ in some important respects from the April encoder -- we have no way of knowing this, since NBS and CBS have declined to give us access to the specifications of the CBS filter actually submitted to NBS for study. Accordingly, many of these selections may not fully illustrate the audible degradation caused by the version actually given to NBS. But the wide musical variety of the selections listed here does indicate that, should the characteristics of the NBS encoder be made known, appropriate musical selections of equivalent variety, probably from different passages of the discs listed here, could also be chosen to "match" its characteristics and thereby most clearly illustrate the distortion...
>
> "The selections listed below were chosen to illustrate a principal sonic liability of any filter fulfilling the requirements of the proposed legislation: removal of significant and audible fundamentals or harmonics from a musical signal. Other sonic degradations from an encoder are possible, including a highly colored noise spectrum and encoder sub-stage overload. Knowledge of the electronic characteristics of the actual CBS filter will allow the NBS consultant(s) to select music suitable for testing these aspects in listening tests.
>
> "All selections are on CD. Although prerecorded cassettes and LPs would equally be suitable for inclusion (and could also be encoded under the proposed legislation), CDs were chosen for their lower noise level, their much better playback repeatability with different pressings and machines, and the ability to perform rapid phrase repeats.

"These selections are believed to be of suitable characteristics and of sufficient duration to reveal the audible degradation caused by sharp band-reject filter meeting the requirements of the proposed legislation and of sufficient bandwidth and depth to allow the design of a reliable notch-detection scanner circuit. This is not to imply, however, that other portions of this, or other, material, is entirely unaffected by such filtering...

"<u>Recommended Selections</u>

"In the list below, the first number for each selection (e.g., "5.") refers to an entire CD. The numbers with capital letters (e.g., "5A, 5B") refer to particular tracks within that CD. This information is followed by the track number and time reading as indicated by the readout on the CD player. The particular material referenced is then further described, occasionally with further notations of "obvious" or "very obvious" (as tested by Mr. Ranada with his filter). The selections are not in any particular order."

1.	Streisand: The Broadway Album. CBS CK-40092. Pressing made in Japan Matrix No. 1A7 73		
1-A	Track 4	1:42 - 1:46	note on word "time"
1-B	Track 9	4:10 - 4:14	note on word "dies"
1-C	Track 10	1:20 - 1:22	note on word "man"
1-D	Track 11	2:11 - 2:16	note on word "friend"
1-E	Track 12	2:51 - 2:53	note on word "time" **very obvious
1-F		2:56 - 3:00	note on syllable "-day"
1-G		3:00 - 3:03	note on syllable "-where"
1-H		3:58 - 4:03	note on syllable "-how"
1-I		4:04 - 4:07	note on syllable "-day"
2.	Prokofiev: Alexander Nevsky. London 410 164-2.		
2-A	Track 7	4:16 - 4:31	piccolo note. NB: note disappears for a moment at 4:25 in unfiltered version **very obvious
3.	Joan Sutherland: Bel Canto Arias. London 417 253-2.		
3-A	Track 4	5:44 - 5:48	soprano high note
4.	Villa-Lobos: Bachianas Brasileiras No. 5. London 411 730-2.		
4-B	Track 14	0:29 - 0:33	soprano note
4-B		1:22 - 1:24	soprano note
5.	Beethovan: Symphony NO. 5. Telarc CD-80060.		
5-A	Track 2	5:16 - 5:20	unison E-flat, full orchestra
5-B	Track 3	11:26 - 11:29	E-flat in violins
6.	Beethoven: Symphony No. 6. Delos D/CD-3017		
6-A	Track 2	11:03 - 11:07	B-flat note in violins

7.	Mozart: Sinfonia Concertante. Deutsche Grammophon 413 461-2.		
7-A	Track 1	1:54 - 2:01	syncopated chords
7-B	"	2:23 - 2:26	
7-C	"	8:07 - 8:11	violin note
	Mozart: Violin Concerto No. 1		
7-D	Track 5	2:06 - 2:20	violin note
7-E	"	5:51 - 6:06	"
7-F	"	9:08 - 9:11	"
8.	Sibelius: Violin Concerto. RCA RCD1-7019.		
8-A	Track 1	5:52 - 5:54	high B-flat in violin
	Prokofiev: Violin Concerto No. 2		
8-8	Track 5	1:21 - 1:23	violin note
9.	Copland: Fanfare for the Common Man Pro Arte CDD-102		
9-A	Track 5	2:18 - 2:20	brass chord
	Barber: Adagio for Strings		
9-B	Track 7	6:55 - 7:07	string chord
10.	Barber: Adagio for Strings. Telarc CD-80059		
10-A	Track 4	4:57 - 5:06	string chord (same as 9-B)
11.	Stravinsky: Petrouchka. CBS MK-37271.		
11-A	Track 1	35:45 - 35:48	trumpet dissonance
12.	Carlos: Switched on Bach. CBS MK 7194.		
12-A	Track 7	0:19 - 0:22	
13.	Beethoven: Violin Concerto. London 400 048-2.		
13-A	Track 2	6:27 - 6:28	violin D-sharp
14.	Beethoven: Symphony No. 9. Angel CDC 7 47189 2.		
14-A	Track 1	0:48 - 0:52	orchestral chord
15.	Beethoven: Symphony No. 9. Denon CD-7021.		
15-A	Track 1	0:41 - 0:48	orchestral chord (same as 14-A)
16.	Stravinsky: L'Histoire du soldat. Delos D/CD 3021		
16-A	Track 8	0:18 - 0:22	chord
17.	Elgar: Violin Concerto. Deutsche Grammophon.		
17-A	Track 2	7:13 - 7:22	B-flat in violin
18.	The unforgettable Glenn Miller. RCA PCD1-5459.		
18-A	Track 4	0:22 - 0:23	telephone bell in "Pennsylvania 6-5-Thousand"
18-B	"	0:35 - 0:36	"
18-C	"	1:01 - 1:02	"
18-D	"	2:13 - 2:14	"
18-E	Track 12	0:09 - 0:22	trumpet notes

18-F Track 13 3:21 - 3:23 last chord of "Serenade in Blue"

19. Debussy: La Mer. Angel/[MI CDC 7 47028 2
19-A Track 1 2:08 - 2:12 violin chord

20. Debussy: La Mar. CBS MYK 37261
20-A Track 1 3:44 - 3:46 violin chord (same as 19-A)
**very obvious

21. Debussy: La Mer. Telarc CD-80071
21-A Track 1 3:33 - 3:36 violin chord (same as19-A)
21-B Track 2 0:45 -0:46 percussion notes

22. Kodaly: Hary Janos Suite. Angel/EMI CDC 7 47109 2
22-A Track 2 0:12 - 0:13 top piccolo note disappears
**very obvious

23. Kodaly: Hary Janos Suite. CBS MYK 38527
23-A Track 2 0:12 - 0:14 top piccolo note disappears (same as 22-A)
23-B Track 2 2:06 - 2:07 piccolo notes disappear
23-C Track 6 2:51 - 2:54 "
Prokofiev: Lt. Kije Suite
23-D Track 10 2:33 - 2:25 "

24. Sousa: Marches. Kem Disc 1004
24-A Track 6 2:39 - 2:40 piccolo notes disappear with filtering during "Anchor & Star" March
24-B Track 7 1:53 - 1:56 "Stars and Stripes Forever"
24-C " 2:48 - 2:50 top piccolo note disappears with filtering
**very obvious

25. Messiaen: Turangalila Symphony. CBS M2K 42271. Disc 1.
25-A Track 2 6:20 loud chord
25-B Track 4 0:29 piccolo note altered
25-C " 1:03 "

26. Mahler: Symphony No. 3. CBS M2K 42403. Disc 1.
26-A Track 15 3:12 - 3:17 orchestral chord with piccolo
**very obvious

27. Messiaen: Organ Music. Calliope CAL 9926.
27-A Track 5 3:20 - 3:34 organ chord **obvious
27-B Track 8 5:17 - 5:37 " **obvious

28. Sousa: Stars and STripes Forever. Telarc CD-80099
28-A Track 11 2:56 - 3:00 top piccolo note disappears (same as 24-C)
**obvious

29. Beethoven: Symphony No. 9. Deutsche Grammophon 410 987-2
29-A Track 5 3:16 - 3:23 chord with chorus "vor Gott"

30. Debussy: Preludes Book I No. 5. Calliope CAL 9831.
30-A Track 5 2:36 - 2:39 top piano note in passage
**very obvious

31. Debussy: Preludes. Angel/EMI CDS 7 47608 8. Disc 1.
31-A Track 5 3:14 top piano note disappears with filtering
**very obvious
31-B " 3:16 - 3:17 same as 30-A
31-C Track 11 0:20 0:22 trilled piano chord

32. Copland: Fanfare for the Common Man. Telarc CD-80078.
32-A Track 1 1:56 - 1:58 brass chord (same as 9-A)

33. Mahler: Symphony 9. CBS M2K-42033. Disc. 1.
33-A Track 1 7:04 - 7:12 String and brass chords

34. Mahler: Symphony 9. CBS M3K-42200. First Movement.
34-A 1st Mvmt. 6:37 - 6:44 Same as 33-A

35. Strauss: Ein Heldenleben. London 414-292-2.
35-A Track 1 3:54 - 3:58 Full orchestra chord
35-B Track 5 5:10 - 5:30 B-flat in upper strings

36. Strauss: Ein Heldenleben. CBC SMCD5036
36-A Track 1 3:00 - 3:04 String chord
36-B Track 1 4:06 - 4:11 Full orchestra chord
36-C Track 2 12:18 - 12:39 Same as 35-B

37. Beethoven: Symphony 2. EMI/ANGEL CDC 7 47698 2.
37-A Track 2 3:19 - 3:25 B-natural in violins
37-B " 7:23 - 7:30 E-natural in violins

38. Shostakovich: Symphony No. 10. RCA 6597-2-RC.
38-A Track 4 10:21 - 10:27 Piccolo runs altered when filtered

39. The Beatles: Rubber Soul. Capitol CDP 7 46440 2.
39-A Track 8 0:06 - 0:08 Syllable "on"
39-B " 0:11 - 0:13 "
39-C " 0:26 - 0:27 "
39-D " 0:48 - 0:50 "
39-E " 0:53 - 0:54 "

40. The Beatles: Revolver. Capital CDP 7 46441 2.
40-A Track 1 1:07 - 1:10 Syllable "oh," backup chorus
40-B Track 4 2:34 - 2:54 Sitar drone
40-C Track 7 2:19 - 2:31 Electric organ

41. The Art of Beverly Sills, Vol. 2, EMI/ANGEL CDC 7 47332 2.
41-A Track 2 9:02 - 9:05 Soprano note
41-B Track 6 5:30 - 5:32 "
41-C Track 9 4:27 - 4:34 "

42. Wagner: Parsifal Prelude. EMI/ANGEL CDC 7 47255 2.
42-A Track B 4:25 - 4:44 Wind chord

43. Verdi: La Traviata. Deutsche Grammophon DG 415 132-2.
43-A Disc 1, Track 11 0:27 - 0:29 violin chord
43-B Disc 2, Track 2 0:00 - 0:39 damage is audible on and off during the track, depending on the chord being played

44. Verdi Otello. RCA RCD2-2951.
44-A Disc 1, Track 10 3:35 - 4:05 Violin Trill
44-B Disc 2, Track 12 0:00 - 0:54 Held violin note
44-C " 2:55 - 3:12
44-D " 4:03 - 4:37 Effect audible depending on chord

45. Beethoven: Symphony No. 9. Telarc CD-80090.
45-A Track 1 0:41 - 0:50 Violin E-flat
45-B " 11:33 - 11:40 "
45-C Track 4 10:44 - 10:50 B-flat in violins

46. Kitaro: My best. Gramavision 18-7016-2.
46-A Track 11 0:27 -0:30
46-B " 0:41 - 0:45
46-C " 2:25 - 2:31

47. Britten: Ceremony of Carols. EMl/ANGEL 7 47709 2.
47-A Track 2 1:14 - 1:17 Chord, boys chorus with harp

48. Bernstein: Chichester Psalms. Deutsche Grammophon 415 965-2.
48-A Track 13 0:01 - 0:04
48-B " 0:07 - 0:09
48-C " 0:32 - 0:36
48-D Track 14 1:16 - 1:19
48-E " 1:34 - 1:35
48-F " 2:45 - 2:48
48-G " 4:51 - 4:54
48-H Track 15 0:04 - 0:06
48-I " 0:19 - 0:22
46-J " 0:48 - 0:50
48-K " 4:16 - 4:20

49. Bernstein: Symphony No. 1. Deutsche Grammophon 415 964-2.
49-A Track 2 0:54 - 0:57
49-B " 1:41 - 1:46
49-C " 5:12 - 5:19

50. Debussy: Nocturnes. Philips 400 023-2.
50-A Track 1 4:26 - 4:43 String chord
50-B " 5:07 - 5:16 "

51. The Art of Beverly Sills, Vol. 1. Angel CDC 47183.
51-A Track 4 6:18 - 6:20

52. Bach: Violin Concerto No. 1. Erato ECD 75358.
52-A Track 5 3:54 - 4:01

53. Strauss: Lieder. Philips 416 298-2.
53-A Track 4 1:09 - 1:12
53-B Track 10 4:27 - 4:30
53-C Track 14 0:58 - 1:00
53-D Track 20 2:43 - 2:46

54. Strauss: Der Rosenkavalier. Angel CDS 7493542. Disc 1.
54-A Track 19 6:55 7:21

55. Wagner: Das Rheingold. Deutsche Grammophon 415 141-2. Disc 1.
55-A Track 1 2:04 - 2:40 Selected notes in violin arpeggio altered
55-B 3:48 - 4:35 "

56. Strauss: Tod und Verklaerung. Chandos CHAN 8533.
56-A Track 4 13:14 - 13:17
56-B " 13:18 - 13:24

57. Strauss: Don Quixote. Angel CDC 747865 2.
57-A Track 2 15:54 - 16:09
57-B " 18:01 - 18:22
57-C " 19:42 - 20:00

58. Scriabin: Poem of Ecstacy. Erato ECD 75360
58-A Track 3 18:37 - 18:42

59. Verdi: Aida. Deutsche Grammophon 410 092-2. Disc 3.
59-A Track 12 0:44 - 0:52 2nd violin B-flat harmonic
59-B " 0:55 - 1:05
59-C " 1:07 - 1:16
59-D " 1:19 - 1:26
59-E " 2:07 - 2:13
59-F " 4:50 - 5:53

60. Holst; The Planets. Deutsche Grammophon Galleria 419 475-2.
60-A Track2 1:35 - 1:54
60-B " 2:40 - 2:56
60-C " 4:55 - 5:02

61. Elgar: Cockaigne. London Jubilee 417719-2.
61-A Track 21 4:25 - 4:33 String chord
61-B " 5:15 - 5:20

62. Bartok: Music for Strings, Percussion and Celesta. London 411 894-2.
62-A Track 9 4:34 - 4:42 string chord

62-B	"	4:46 - 4:55	
62-C	Track 10	1:52 - 2:00	Strings B-flat, left channel
62-D	Track 12	1:18 - 1:23	

63. Bartok: Concerto for Orchestra. London 400 052-2

63-A	Track 1	2:36 - 2:41

64. Sibelius: Symphony No. 7. Bis BIS-311.

64-A	Track 1	16:32 - 16:35
64-B	"	16:36 - 16:39
64-C	"	18:56 - 19:20

65. Bernstein: West Side Story. CBS CK 32603.

65-A	Track 1	1:53 - 1:58
65-B	Track 5	0:44 - 0:47
65-C	"	3:28 - 3:36
65-D	Track 8	2:38 - 2:45

66. Sibelius: Finlandia. Angel Studio CDM-769017-2.

66-A	Track 1	4:54 - 5:41
66-B	"	5:38 - 5:41
66-C	"	6:29 - 6:31
66-D	"	6:50 - 6:52
66-E	"	7:11 - 8:16
66-F	"	8:55 - 9:10

67. Schumann: Symphony No. 3. Deutsche Grammophon 415 358-2.

67-A	Track 1	2:57 - 3:05
67-B	"	6:05 - 6:25
67-C	"	9:05 - 9:24

68. Bruckner: Symphony No. 4. Angel Studio CDM-769006-2.

68-A	Track 1	4:08 - 4:16
68-B	"	20:01-end of movement

Recording Industry Association of America

The 35 selections listed below were sent to NBS on September 29, 1987. Many of these selections were reviewed as to their suitability for inclusion in the listening tests.

Selection	Artist	Label
1. Bartok: Hungarian Pictures	Equale Brass	Nimbus
2. Copland: Rodeo. Hoedown	Atlanta Sym.	Telarc
3. Mendelssohn: Sonata for Piano/ Klavier. 4..	Perahia	CBS/Sony
4. Bach/Gounod: Ave Maria	Shinozaki	Devon
5. Chopin: Piano Concerto No. 2 in F minor	Davidovich	Philips
6. McCartney/Jackson: Say Say Say	McCartney	CBS/Sony
7. Tschaikowsky: Sym. No. 6 Pathetique 3.	L.A. Philar	DG
8. Kabalevsky: con. No. 1 for Cello & Orch. 2	Yo Yo Ma	CBS/M
9. Tschaikowsky: Sym. No. 6	L.A. Philar	DG
10. Sting Spirits in the Material	Police	(AM)
Il. Suite for Flute and Jazz Piano	Rampal	CBS/M
12. Bach: Violin Concerto in E major l.	Mutter/Accardo	EMI
13. Haydn: Horn Concerto No. 2 in D. 2	Wallace/Thomp.	Nimbus
14. J. Taylor: There We Are	Taylor	CBS
15. Rodrigo: Concierto de Aranjuez l.	J. Williams	CBS/M
16. Bach: Concerto for Oboe, Violin & Orch. 3.	Stern/Zukerman	CBS/Sony
17. Schubert: Arpeggiona Sonata. l.	Galway	RCA
18. Momentary Lapse of Reason	Pink Floyd	CBS
19. Schubert: Arpeggione Sonata. 2.	Galway	RCA
20. J. Zawinul: Dream Clock	Weather Report	CBS/Sony
21. Malcolm Arnold: Quintet. l.	Equale Brass	Nimbus
22. Ellington: Rockin'in Rhythm	Weather Report	CBS/Sony
23. Copland: Rodeo. Corral Nocturne	Atlanta Sym.	Telarc
24. Beethoven Symphony No. 9 (Choral)	Walter	CBS/M
25. E. Jones: The Things I Used to Do	S.R. Vaughn	Epic
26. Haydn: Trumpet Concerto in E flat 3.	Wallace/Thomp	Nimbus
27. Schubert: Serenade	Galway	RCA
28. Mendellssohn: Prel. and Fugue Op. 35, No.l	Perahia	CBS/M
29. George Strait #7	Strait	MCA
30. Vivaldi Four Seasons	Pinnock	Archiv
31. Some	Belouis	EMI
32. In Square Circle	Wonder	Motown
33. Billie Holiday Story	Various	CBS
34. Liszt Transcendental Studies	Horowitz	Telarc
35. Turbo	Judas Priest	CBS

APPENDIX E

Description of Recorded Material for Listening Tests

Appendix E: Description of Recorded Material for Listening Tests

The musical selections that were included in the serial and parallel listening tests are identified in table E-1, along with a description of the observed effects that led to their inclusion in the study. The first column of table E-1 includes an item number, corresponding to the order of presentation for the first 24 trials in the serial listening study. An asterisk after the item number indicates that the particular selection was suggested by the HRRC while a plus sign means that the RIAA suggested that compact disc. Immediately below the item number is given the Serial Number of the Encoder (see section 2.1) used to encode the musical selections for both the serial and the parallel listening tests. The second column provides the identification of the compact disc. The third column indicates, as applicable, the composer, the composition, and the performer. The fourth column provides a brief identification of the selection for use in identifying it in the color spectrographs given in section 4.4 and in this appendix, as well as in the tables in section 4.5 and Appendix C. Column 5 indicates the particular track on the compact disk, as well as the time period from that track -- the time interval, in minutes and seconds, given just below the Track Number corresponds to the short selections use in the serial listening tests while the second time period, if any, that is enclosed in square brackets designates the portion of the track that was used for the parallel listening tests.

The final column of table E-1 indicates which, if any, of the six hypothesized effects discussed in section 4.4 were believed to be potentially audible for that particular selection. The text across the bottom of the block for each selection describes the effects that were observed by NBS staff for that selection. As mentioned in section 4.4, NBS staff utilized physical measurements and real-time spectrographs to assist in identifying these particular selections and the potentially audible effects of the encoder. Extensive listening by NBS staff was then performed in order to determine whether the effects of the Encoder could be heard and to relate the audible effects, if any, to the hypothesized effects.

After the selections were recorded to digital audio tape, in both direct and encoded form, these tapes were used as the source material for the color spectrographs shown on the cover of this publication, in sections 2.4 and 4.4, and in this appendix. These spectrographs were obtained by storing the sum of the left and right channels for each musical sample on a disc using the "throughput" mode of an HP 3562A analyzer. The input range of the analyzer was set so the peak level of each sample was just below overload. Then, the data were analyzed using the Fast Fourier Transform capability of the analyzer to obtain the power spectrum of each record, each record being 128 ms long. Thus, there are approximately 7.8 records per second in the spectrographs. The analyzer bandwidth was 0 to 6.25 kHz, with 801 frequency bands in the FFT analysis. Each odd-numbered frequency band from the analysis was averaged with the next-higher even-numbered band, reducing the 801 bands to 400. Since only frequencies up to 6 kHz are displayed, each record contains 384 separate frequency bands, spaced 15.62 Hz apart.

Each record from the Analyzer was read into an HP 330 computer, and processed to transfer the complete spectrograph onto an HP PaintJet printer. The dyna-

mic range of compact discs is on the order of 90 dB, so this range was chosen for display on the spectrographs, with a range of 10 dB for each of 9 colors. Because the dynamic range of the analyzer is 80 dB, the full 90 dB cannot be displayed for any one selection. However, the input range was set to make optimum use of the analyzer for each selection, so 90 dB is displayed over all samples.

Examples of the many spectrographs that were made during the testing are illustrated on the cover of this report and in section 4.4, figures 32 and 33.

Table E-1 Musical Selections for Listening Tests

Item & Encoder	CD identification	Title and artist	Short I. D.	Track no. and time [extended time]	hypothe-sized effect(s)
1. * CCE-006	EMI CDC 7470282	Debussy: La Mer Previn, London symphony orch.	Debussy	1 2:09-2:12	2
A strong harmonic of the violins is removed.					
2. * CCE-012	Telarc CD80059	Barber:Adagio for Strings Slatkin, St. Louis Symph.	Barber	4 4:57-5:01	2,3,5
The chord which begins at 4:57 has the upper part of the spread of a harmonic removed. At 4:59 a new chord begins while the original chord continues. At 5:00, the sound naturally brightens slightly, although no new note is started. This brightening causes the encoding notch filter to be switched out, causing an increase in the level in the vicinity of 3840 Hz.					
3. + CCE-012	CBS MK42014	Beethoven:9th symphony Bruno Walter, Columbia Symph. Orch.	Beethoven chord	1 0:38-0:42	2,3
The upper portion of the harmonic of an orchestral chord is removed.					
4. * CCE-006	DG 41098-2	Beethoven:9th symphony Karajan:Berlin Philharmonic	Beethoven chorus	5 3:16-3:23	2?
The spectrum of a full chorus and orchestra had minor components in the encoding region, which were removed.					
5. CCE-012	Denon 33C 39-7441	Bizet-Sarasete Carmen Fantasy Paganini Ensem.	Bizet	23 0:15-0:18 [0:15-1:10]	2,6
The lower half of the spread of a harmonic of a solo violin is removed.					

[Table E-1, cont'd]

6. CCE-012	Denon 33C 39-7441	Billy Joel: Just the Way You Are Nancy Wilson	Wilson	14 0:27	2
Syllable "Nev" of word "Never" has harmonic removed.					
7. * CCE-012	CBS CK40092	Bernstein: Somewhere Barb. Streisand	Streisand	11 2:56-3:01 [2:56-4:15]	2,5
Two different words have a strong harmonic removed.					
8. + CCE-012	CBS MK42014	Beethoven:9th Symphony Bruno Walter, Columbia symph.	Beethoven Triangle	4 10:56-11:01 [10:45 -11:30]	1,2
One of the partials of a triangle is completely removed.					
9. CCE-012	CBS MK42314	All the Things You Are:Kern/ Hammerstein Maur. McGovern	McGovern	2 3:03-3:06	2,5
A strong harmonic of the voice slides upward through encoding notch frequency at beginning of word "are", with attendant loss of harmonic.					
10. + CCE-006	ARC/Columbia CK36793	Dream Clock Weather Report	Weather	2 5:39	1,2
A prominent partial of a cymbal tap was reduced in level.					
11. * CCE-012	Telarc CD80078	Copland: Fanfare for the Common Man Lane, Atlanta Symph.	Copland	1 1:55 [0:53-2:03]	2,5
The brass instruments have a strong harmonic removed.					
12. CCE-006	DG 410025-2	Bernstein:Sym- phonic Dances Bernstein, Los Angeles Phil.	Berstein	11 0:41 [0:00-1:02]	2,3
A strong harmonic of high-pitched violins is removed.					

[Table E-1, cont'd]

13. + CCE-006	Columbia CK40599	Sorrow Pink Floyd	Pink Floyd	10 0:52-1:02	2,5
A complex overloaded guitar sound has harmonic components removed.					
14. CCE-006		Synthesizer No. 1	S1		
A two-note chord of F_7 and B_7. The level of the B_7 is reduced.					
15. * CCE-012	London 410164-2	Prokofiev: Alex. Nevsky Chailly, Cleveland Symph.	Prokofiev	7 4:16-4:21 [4:00-end]	5
The Encoder is switched in and out on a sustained chord, with full orchestra and chorus, erratically reducing a very audible component at the lower edge of the encoding notch frequency region.					
16. CCE-012		Synthesizer No. 2	S2		
A trumpet fanfare in the key of D has harmonic components removed.					
17. CCE-012	ECM 1155	Pat Metheny Amer. Garage	Metheny	4 0:00-0:02 [0:00-0:29]	2,4
Components in the encoding notch frequency region are removed from alternating drum and cymbals. The cymbals originally had a minor formant at 3840 Hz.					
18. CCE-012		Synthesizer No. 3	S3		
This is a sustained E-flat$_6$, with several other notes played consecutively.					
19. CCE-006		Synthesizer No. 4	S4		
This is a sustained chord of B_4, E_5, and F-sharp$_5$, with several other notes being played consecutively. A partial of the E_5 is reduced.					

[Table E-1, cont'd]

20. * CCE-006	CBS M2K42271, disc 1	Messiaen: Turangalila Salonen, Philharmonia Orchestra	Messiaen	4 1:03 [0:27-1:05]	1,2
The first part of a piccolo note, whose fundamental frequency falls in the passband of the encoding filter, is removed.					
21. CCE-012	Archiv 415 112-2	Bach: 3-part Invention No. 2, Gilbert harpsichord	Bach	17 0:23-0:27 [0:00-0:31]	2,4
A trill has a pronounced harmonic for each of the two notes of the trill removed.					
22. CCE-006	Telarc CD80085	Respighi: Fountains of Rome, Lane Atlanta Symph.	Respighi	13 1:58 [1:50-2:35]	1,2
A strong partial of a bell is markedly reduced in level.					
23. CCE-006	L'oiseau-Lyre 411858-2, disc 1	Handel:Messiah Nelson, Soprano Hogwood, Acad. Ancient Mus.	Handel	11 1:58	6
A soprano trill has the center frequencies of its harmonics removed.					
24. CCE-006	Telarc CD8052	Ravel: Daphnis et Chloe Slatkin, St. Louis Symph.	Ravel	3 2:07-2:10	2
The first and, especially, the third of a sequence of 3 notes on a clarinet have a prominent harmonic removed.					

* HRRC list
\+ RIAA list

APPENDIX F

Preparation of Recorded Material

APPENDIX F: PREPARATION OF RECORDED MATERIAL

Equipment Used

The equipment used for making all of the recorded materials presented in both the serial and parallel listening studies consisted of a Nagra Model IV-S analog tape recorder, a Denon Model DCD-3300 compact disc player, a Kurzweil Model 250 synthesizer, components of a compact disc mastering system manufactured by Sony, a Sony Model PCM-3324 multi-channel digital audio tape recorder with a Model RM-3310 remote control, and a digital interface box. The analog tape recorder was used to make the original recordings of the announcements and the tone that were a part of presenting the material to the listening subjects. The compact disc player was used to transfer, from compact disc to digital audio tape, the compact-disc material that was evaluated. The synthesizer was used to create and store the synthesized material that was evaluated.

The Sony disc mastering system components included two Model DMR-4000 digital master recorders, two Model PCM-1630 digital audio processors, one Model DAE-1100A digital audio editor, one Model DABK-1630 (RAR-1) printed circuit board, one set of Model DABK-1631 digital input/output printed circuit boards (only the input board was utilized), and several Model D-3/4-75 cassette tapes. A more detailed description of the characteristics of these components and the components of the multi-channel recording system follows.

The switch settings and any modifications of the printed circuit boards used in the two audio processors were as follows. The Model DA-15 digital-to-analog printed circuit boards were modified by replacing the low-pass filters AFL101 and AFL201 with Apogee Electronics Corporation Model 944-G anti-aliasing low-pass filters for both audio processors. The Model AD-23 analog-to-digital printed circuit board installed in the audio processor used to make the recordings was modified by replacing the low-pass filters AFL101 and AFL201 with Apogee Electronics Corporation Model 944-G anti-aliasing low-pass filters. (Note: all use of these printed circuit boards including recording, playback, testing, and calibration was with the Apogee filters installed in them.) The setting of the switches on the printed circuit boards AD-23, ENC-2, SIF-1 DEC-15, and MT-16, in both audio processors were set to the factory preset positions as given in the 4th or 5th revision of the 1st edition of the PCM-1630 operation and maintenance manual with the exception of the EMP (emphasis) switch on the AD-23 board and the REC MUTE (record muting) switch on the ENC-2 board which were normally set to OFF except as needed. The switches on the optional DABK-1630 (RAR-1) printed circuit board installed in the audio processor used in making the recordings were set to the factory preset positions as given in the 3rd revision of the 1st edition of the DBAK-1630 operation and maintenance manual. The muting time was set for one second on this board. (These settings are included in the interest of completeness since the board was installed in the audio processor during the recordings. However, the read-after-read function, available through this board was never used since the PB MODE selector switch on the audio processor was always in the A position). The DI-5 digital interface printed circuit board was the only one of the two boards included in the DBAK-1631 set of digital I/O boards

that was used. The switches on this input board were set as follows: DIP SW1 - switch 1 was OFF, i.e., set for Sony/Philips consumer formatted data, with switches 2 through 8 ON; the manual emphasis switch (SW-2) was OFF; and the DI SYNC switch (SW3) was ON (see the 2nd revision of the 1st edition of the DBAK-1630 operation and maintenance manual). Only one of the two audio processors was used to make all of the digital recordings, the other was used to play back the recorded materials during the serial presentation.

The switches of the DAE-1100A were set for serial data and word sync and a sampling frequency of 44.1 kHz (DROP FRAME OFF). The switches of the keyboard were set as needed. Player A was the only player connection block that was utilized. The input select switch (IN SEL) in this block was in position D (digital input signal) when the compact disc player output signal was used and position V (composite digital input signal) when the digital master recorder output signal connected to PLAYER-A block was used.

The switches of the DMR-4000 digital master recorder internal printed circuit boards DM-49, SY-37/SY37A, and TC-38 were set to the factory preset positions as given in the 5th revision of the 1st edition of the DMR-4000 operation and maintenance manual. These recorders were always used with the METER switch in the R/P position, the AUX LIMITER switch OFF, the RAW OUT selector in the MAIN position, and the SKEW control in the fixed position. The tracking control was normally in the fixed position except as needed. All the recordings made using these machines were done with the recorder and/or player in the PB/EE mode of operation. Normally the synchronization of these machines was done by the digital editor in the editing process during which time the AUX CH2 selector was in the AUDIO position. However, when it was desired to make a copy of a tape synchronously, with the time code recorded at the same time as the audio, the AUX CH2 selector switches of the player and recorder were in the TIME CODE position with the TIME CODE selector switch in the REC/RUN position in the case of the player and the REGEN position in the case of the recorder (for these tapes the time code output of the player was connected directly to the time code input of the recorder). In all cases, the mastering recorders were used with the 75 ohm composite digital input termination ON.

The switches of the PCM-3324 multi-channel digital recorder printed circuit boards ENC-1, ENC-2, ENC-3, DET-1, DET-2, DET-3, DIN-1, DIN-2, DIN-3, A/D-1 through A/D-6, TBC, DCC, CLK, EDTB, MCUP, EIO, LOC, CTL, SRVO, and SYS boards were set to the factory preset positions as given in volume 1 of the 5th revision of the 3rd edition of the PCM-3324 operation and maintenance manual except as follows. The double indication/calibration selector switch was set to DOUBLE IND, the emphasis switches on the A/D boards were set as necessary, the XFADE (cross fade time) control on the MCPU board was set in position F, which corresponds to a nominal crossfade time of 1.45 ms when using a sampling frequency of 44.1 kHz, and switch 5 of DIP SW4 on the MCPU board was set to OFF thus defeating the AUTO REC function of analog inputs A1 and A2. (Note: it was not verified that the adjustment of the tension controls on the SRVO board corresponded to the factory preset adjustment values.) The 75-ohm word sync termination switch located on the PCM-3324 was switched ON and the normal/external word selector on the VCLK (video clock) board was set in the EXT WRD position. The sampling rate used in the recordings was 44.1 kHz and the INPUT SELECT and SYNC CLOCK switches on the toggle switch (TSW) block of

the PCM-3324 were set in the DIGITAL and AUTO positions, respectively. During the parallel listening tests, the MASTER SAFE switch located on the TSW block was set in the ON position. The Model RM-3310 remote control unit was operated without the audio control unit connected to the system control unit and the selector switches on the rear of the unit were set to provide 100 cue registers.

A custom (not mass produced) digital interface (converter) was used to interface the balanced outputs and inputs of the PCM-3324 with the unbalanced outputs and inputs of the PCM-1630 audio processor. The circuits and components implemented in the design of this unit were based on the circuit descriptions for conversion between balanced RS-422 and unbalanced TTL signals given in the August 1987 revision of the PCM-3324 product information manual. The interface was tested using a function generator and oscilloscope prior to using it to make recordings for the listening studies.

Two rolls (one for test signals and one for program material) of Sony Model D-1/2-1460 digital audio master tape and a Sony Model RH-14DA empty reel were used in making and playing the multi-channel recordings.

Serial Presentation

Copying, Encoding, and Editing Procedures

Figures F-1 through F-3 illustrate the functional arrangement of the equipment used in preparing the material for serial presentation with the exception of the recording of the synthesized selections, announcements, and 3840 Hz tone. Unless otherwise noted, all cassettes used in the recording process were "pre-striped" (time code was recorded on the cassettes prior to recording audio) using the DAE-1100A digital audio editor. The digital sampling frequency used in making all recordings was 44.1 kHz.

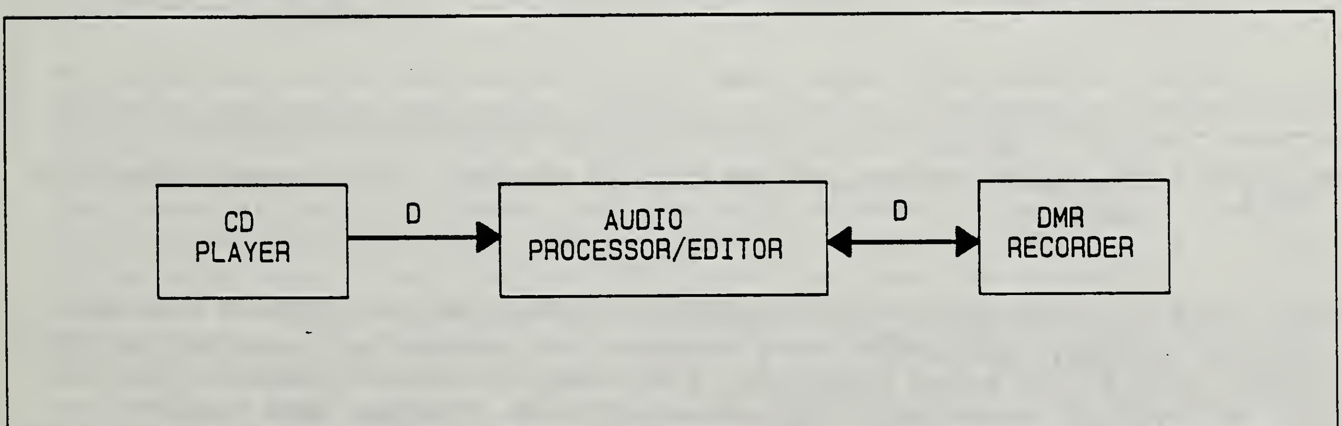


Figure F-1 Equipment setup used for transfer from compact disc to digital audio tape. ("D" denotes digitized audio)

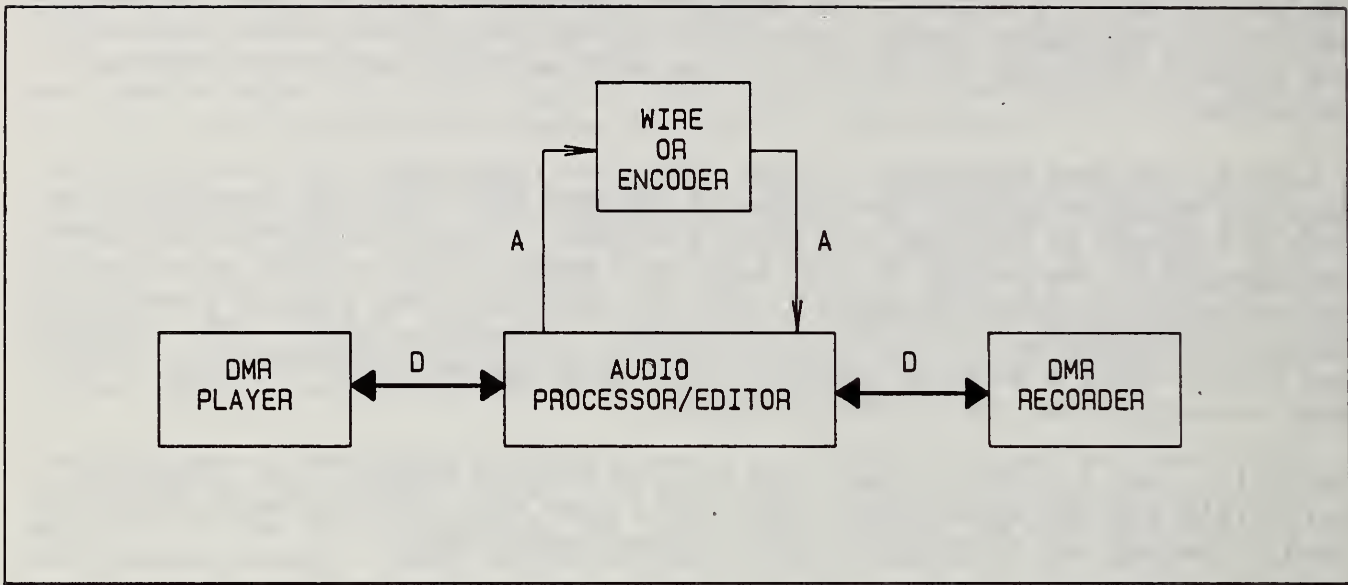


Figure F-2 Equipment setup used for making direct and encoded digital tapes. ("D" denotes digitized audio; "A" denotes 2-channel analog audio)

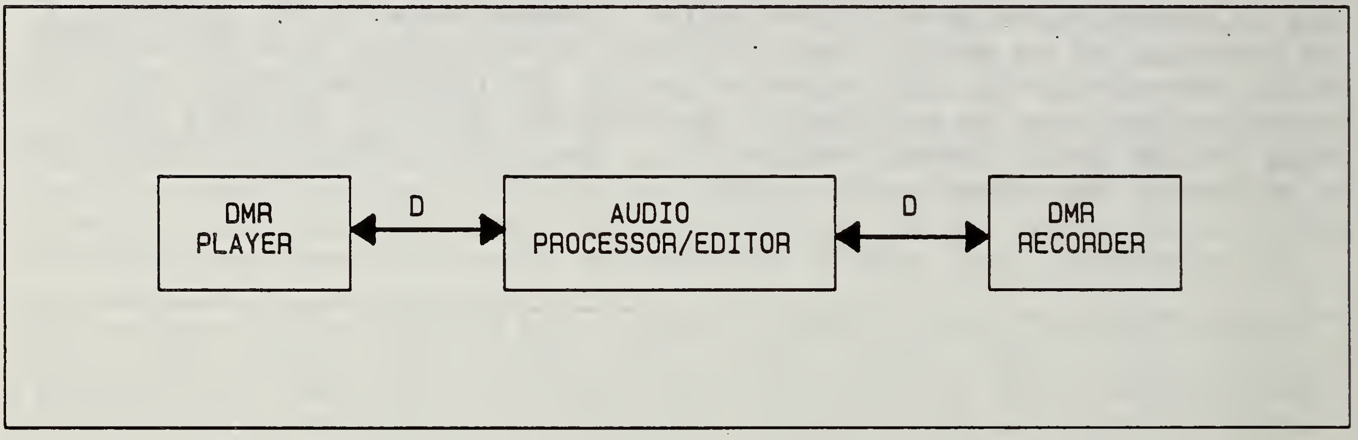


Figure F-3 Equipment setup used for digital editing. ("D" denotes digitized audio)

When recording the synthesized selections, the balanced outputs of the synthesizer (Kurzweil Model 250) were connected to the analog inputs of the AD-23 (analog-to-digital) board installed in the PCM-1630 audio processor used to make the digital recordings. The selections were recorded onto cassette through the DAE-1100A (digital audio editor) using a DMR-4000 (digital master recorder). The sequencer in the Model 250 was used to playback prerecorded sequences, or selections during these recordings. The announcements and tone were recorded, edited and leadered using an analog tape recorder (Nagra IV-S in the case of recording the material). The balanced outputs of the Nagra

were connected to the PCM-1630 and the final analog tape containing the announcements and tone was copied onto cassette using the DAE-1100A and a DMR-4000 in the same fashion as the synthesized recordings. The material that originated on compact disc was recorded onto cassette by connecting coaxial digital output No. 1 of the DCD-3300 (compact disc player) to the DI-5 (digital input) board installed in the PCM-1630 used to make the recordings (see figure F-1). A twisted pair of shielded wires was used to connect the DCD-3300 to the PCM-1630 with pin 3 of the XLR input connector on the PCM-1630 connected to the center conductor of the coaxial output of the DCD-3300 and pin 2 of the XLR input connector on the PCM-1630 connected to the outer conductor of the coaxial output (chassis) of the DCD-3300. The shield was connected to pin 1 of the XLR input connector terminating at the PCM-1630 but not to the shell of the XLR input connector or to the DCD-3300. The chassis of the DCD-3300 was left isolated from ground by the input stage of the DI-5 board. Of the shielding/grounding schemes measured, this connection arrangement provided the fastest rise time in the waveform of the data stream and resulted in a peak-to-peak voltage greater than 0.5 volts in the digital waveform. (Note: the DI-5 digital input board was set specifically to decode Sony/Philips consumer formatted data, not AES/EBU formatted data. Since the two formats are similar but not identical, there may be some commercially available equipment that is designed to decode AES/EBU formatted data that also accepts Sony/Philips consumer formatted data but not without some amount of error occurring in the decoding process, e.g., incorrect recognition of emphasis information.) The selections were recorded onto cassette through the DAE-1100A (digital audio editor) using a DMR-4000 (digital master recorder). The GAIN OFFSET fader and switches on the keyboard of the DAE-1100A were used to create the beginning and end of each of the twenty test selections plus four practice selections taken from compact disc. There was no overall gain offset introduced into the recorded material during this recording process, i.e., the gain offset was set for 0 decibels as shown by the illumination of the LED adjacent to the gain offset fader, indicating that the gain offset function was not in use between fading or switching the material in and out. These recordings, plus the recordings of the synthesizer, comprised a master cassette of the twenty-four selections used in the serial study plus the four practice selections.

The selection master was copied three times: once with shielded cables directly connecting the outputs (analog) of the DA-15 digital-to-analog board of the PCM-1630 to the inputs (analog) of the AD-23 analog-to-digital board of the PCM-1630, once with Encoder 006 between the outputs of the DA-15 board and the inputs of the AD-23 board, and once with Encoder 012 between the outputs of the DA-15 board and the inputs of the AD-23 board (see figure F-2). These copies were made through the DAE-1100A with the time code output of the player connected directly to the time code input of the recorder and 0-dB gain offset in the editor. The recorder and player were operated synchronously with the AUX CH2 selector switches of the player and recorder in the TIME CODE position and the TIME CODE selector switches in the REC/RUN position in the case of the player and the REGEN position in the case of the recorder. The time code on the selection master was recorded onto the three copies at the same time that the audio material was recorded so that the audio material being recorded had the same relative position with respect to the time code track for each of the three copies. The rms voltage levels at the output of the two channels of the

PCM-1630 were adjusted to +20.0 dB relative to $\sqrt{0.6}$ V rms ($\sqrt{0.6}$ V rms corresponds to 0 dBm) using the 1-kHz, 0-dB digital reference level from track 1 of the CBS CD-1 test disc recorded onto the selection master cassette during the same time period that the selection material was copied from compact disc. (Note: the term "dBm" as it has been used in this appendix denotes voltage level referenced to 1 milliwatt of average electrical power dissipated in a purely resistive 600 ohm electrical load. However, the use of the term does not imply that the voltages that were measured were terminated by a purely resistive load, or that they were derived from measurements of power.) The drift observed in the PCM-1630 output levels over the duration of the time period that the three copies were made was less than 0.1 dB. Since some of the recorded material had peak levels equal to 0 dB (maximum digital level) and because the gain of the Encoder was slightly greater than unity and the AD-23 input stages exhibited a slight rise in response at low-frequencies relative to mid-band, the gains of the analog input stages of the AD-23 analog-to-digital board were set for -1.0 dB at 1 kHz to prevent clipping the signals during analog-to-digital conversion. (Note: since the distortion of the input stages of the AD-23 board was measured and found to be a function of input level (see the following section on equipment calibration), the rms levels of the selections were measured using a Bruel and Kjaer Model 2305 graphic level recorder and found to be at least eight decibels below the maximum peak recorded levels. Thus, at no time was the rms level of the voltage at the input stages of the AD-23 board greater than +12 dB relative to $\sqrt{0.6}$ V rms.) The lengths of the cables connecting the inputs and outputs of the PCM-1630 during the "direct" recording were approximately 1 m. These cables were connected between the outputs of the PCM-1630 and the inputs of the Encoder during the "encoded" recordings. The lengths of the cables connecting the outputs of the Encoder to the inputs of the PCM-1630 during the "encoded" recordings were approximately 1.5 m. In all cases, the shields of the cables were connected to pin 1, but not the shell, of the connectors at each end of the cables. The PCM-1630 and the Encoder were grounded at their respective AC power connectors using the same AC power circuit (see the following section for the differences in the "direct" and "encoded" noise spectra at the PCM-1630 inputs with this grounding arrangement). The Encoder was set for fixed gain and direct (non-transformer) coupled outputs. The Encoder was switched IN throughout the duration of each of the encoded copies. Digital emphasis was added to the Weather and Bernstein selections at this stage of the recording process by activating the emphasis circuitry in the input section of the PCM-1630 with the EMP (emphasis) switch on the AD-23 analog-to-digital board during each of the copies of these selections. (Note: the operation of the EMP switch produced switching transients, in each of the copies, which were removed in the next stage of the recording process. The amount of time that the emphasis circuit was active before and after the recorded material began and ended was also made the same for the emphasized selections at this time.) The recordings were monitored using headphones connected to the PCM-1630. These recordings, one direct and two encoded, were then copied and edited using the DAE-1100A.

The three copies, "Direct," "E006," and "E012," were assembled into four tapes: "DD," "DE," "ED," and "EE" using the DAE-1100A editor (see figure F-3). These four tapes consisted of the four possible combinations of the pairs of each of the twenty-four test selections repeated three times with tape DD

containing the direct/direct pairs, etc. In addition, tape DD contained the four practice pairs presented ahead of the actual test pairs. The tapes were assembled in the following manner. Because the material recorded on the three copies of the selection master had the same relative position with respect to time code, only one "play" or start point on the playback machine (EDIT-IN point in DAE-1100A terminology) was needed for any given selection, so that DD, DE, ED, and EE were edited using a total of twenty-four play points for the test selections plus four play points for the practice selections. The play points were defined 0.5 seconds before the first audible point of each of the selections. The first "record" or start point on the recording machine (EDIT-OUT point in DAE-1100A terminology) was arbitrarily defined for numerical convenience. The second record point was defined 0.5 seconds after the last audible point of each of the selections, so that after the second electronic edit was complete the first presentation of the selection pair being edited was complete with a one-second pause existing between the members of the first pair in the presentation (0.5 seconds for the play point plus 0.5 seconds for the record point). The third record point was defined by taking the difference between the first and second record points to find the length of each pair member and adding this time plus two seconds to the time defining the second record point. This resulted in a three-second pause between the end of the first and the beginning of the second presentation of the pair being edited (0.5 seconds for the play point plus 2.5 seconds for the record point). With the length of the pair member established, the remaining record points were defined so that there was a one-second pause between pair members and a three-second pause between pair presentations. For each set of play and record (EDIT-IN and EDIT-OUT) points, an electronic edit (AUTO EDIT) was performed four times using the four tapes DD, DE, ED, and EE for recording and two of the three tapes DIRECT, and E006 or E012 for playback (except for the practice pairs which were presented in only one combination of pairs). Because the editor synchronized the operation of the player and recorder during each electronic edit and the edit points were, by definition, the same for each of the four tapes, each selection was presented in the same way (the same pauses between pairs and pair members, beginning and ending points, etc.) for each of the four presentations of possible combinations of pair members. During this stage of the recording process, relative levels of the individual selections were adjusted using the GAIN OFFSET fader in the DAE-1100A. The level of the Streisand and Beethoven chorus selections were reduced by approximately 8 dB; the level of the Prokokiev selection was reduced approximately 7 dB; the levels of the Bach, Pink Floyd, and S2 selections were each reduced by approximately 1 dB; the level of the Bernstein selection was raised approximately 6 dB; and the level of the Weather selection was raised approximately 4 dB. Once the gain offset relating to each of these selections had been established, it remained unchanged during the recording of each of the selections onto the four tapes. These relative level adjustments prevented excessively loud or soft playback of individual selections during the serial presentations.

The four tapes DD, DE, ED, and EE were then edited, with the tape containing the announcements, into two final tapes for presentation, S1 and S2 (see figure F-3). The ordering and pairing of the material on these tapes corresponded to the randomization established by the NBS Statistical Engineering Division for Tests 1 through 96. During the editing process, the record

levels of the announcements were adjusted slightly as needed using the GAIN OFFSET fader so that the level of the announcements corresponded more to the level of the material following or preceding each announcement i.e., relatively loud or soft (the GAIN OFFSET fader was set to 0 dB gain as shown by the illumination of the LED indicator adjacent to the fader indicating that the gain offset function was not in use when recording the selections of test material). The edit points were selected so that there was a 1.5-second pause between the announcements introducing and concluding the test material and the beginning and end of each of the presentations of test material. The beginning and end points of the material were those defined during the editing of tapes DD, DE, ED, and EE i.e., the edit points were defined so that the recorder began recording 0.5 seconds before and after the beginning and end of the test material. The S1 and S2 tapes were checked after completion to verify that the pairing of direct and encoded material for each of the Tests matched that established by the NBS Statistical Engineering Division and that the pairing was consistent throughout the presentation of each Test (the HP 3565S signal processing system and 330 computer mentioned in section 4.2 were used for this analysis).

Equipment Calibration.

The equipment used to calibrate and test components of the recording system and the recording system as it was used to make the tapes for the serial presentation were: (1) a copy of the CBS CD-1 test disc for measuring CD player performance with EIA standard signals, (2) an Audio Precision Model System One-A dual-channel input/output audio test system with a Zenith Model ZWX-248-62 computer, (3) a Fluke Model 8860 digital multimeter with true rms AC voltage measurement capability, and (4) a Tektronix Model 7704 oscilloscope with a Model 7A18 dual trace amplifier and a Model 7B70 time base.

The Audio Precision test system was calibrated by the members of the Electronic Instrumentation and Metrology Group at NBS. Unless noted otherwise, the Audio Precision test system was used without any bandwidth control filters resulting in a nominal measurement bandwidth of 10 Hz to 500 kHz. Thus, the results of noise measurements, or measurements that include noise such as total harmonic distortion plus noise (THD+N), include any ultrasonic noise components present in the measured signal due to the digital sampling frequency of 44.1 kHz. Unless noted otherwise, the 600 ohm input terminations of the Audio Precision were used when making measurements with this unit. The shields of all balanced connections to the Audio Precision were terminated at pin one, but not the shell, of the XLR connectors at the ends of the cables connected to the test system. The outputs of the Audio Precision were set for 50-ohm balanced operation. Only the XLR connectors on the Audio Precision were used in connecting it to the PCM-1630. (Note: the convention used for the assignment of signal high and low to pins 2 and 3 of the XLR connectors are reversed in the case of the Audio Precision test system and the Sony PCM-1630 audio processor. No attempt was made to correct for this 180 degree phase difference in the units when connecting the outputs and inputs of the two instruments since only relative phase measurements were performed using the Audio Precision.) All components of the Audio Precision test system were grounded at their AC power connector when in use.

The Fluke Model 8860 digital multimeter was calibrated using a Fluke Model 5200A precision AC calibrator that in turn was calibrated by the members of the Electronic Instrumentation and Metrology Group at NBS. The Fluke Model 8860 was used to set the gains of the analog input and output stages of the PCM-1630 during the recording of the Direct, E006, and E012 tapes. The nominal AC measurement bandwidth of this instrument is 20 Hz to 300 kHz with a nominal upper frequency limit (-3 dB point) at 1 MHz. The nominal input impedance of the 8860 is 10 megohms with an input capacitance of 70 pF or less. The AC voltage measurements made using the Fluke 8860 were guarded with the shield of the cable connecting the meter to the voltage source connected to pin one, but not the shell, of the XLR connector at the voltage source and to the guard circuitry of the digital multimeter at the Fluke 8860. The guard circuitry was grounded at the voltage source (Fluke 8860 EXT-GD mode of operation). The CBS CD-1 compact disc was copied onto cassette through the DAE-1100A, without gain offset, for future use in testing the recording system. With the compact disc player still connected to the PCM-1630 the outputs of the digital input board were connected directly to the inputs of the DA-15 digital-to-analog board. The balanced analog outputs of the DA-15 board were connected directly to the 600-ohm inputs of the Audio Precision test system and a number of measurements were performed using the test signals on the CBS CD-1 test disc. The output voltages of the DA-15 board were nominally adjusted to 7.75 Vrms (+20 dB re $\sqrt{0.6}$ volts rms, i.e., +20 dBm) using the measuring voltmeter in the Audio Precision and the 1-kHz, 0-dB reference level on track 1 of the test disc. The results of these measurements were used to characterize the performance of both channels of the DA-15 digital-to-analog board in combination with the Denon DCD-3300 as it was connected to the PCM-1630 during the copying of the compact disc material.

A summary of these measurement results, and the test-disc track and index numbers associated with them, follows. The decibel terms in the linearity measurements refer to "0 dB" as the maximum (peak) digital signal level in the same sense as it is used in the description of the test signals on the test disc. The term "1-kHz reference" refers to the voltages present at the analog outputs of the PCM-1630 during the playback of the 1-kHz reference level on track 1 of the CBS test disc. SMPTE refers to the Society of Motion Picture and Television Engineers.

Frequency response without digital emphasis 17 to 19997 Hz: within ± 0.3 dB (Track 6 Index 03 through Track 10 Index 01).

Frequency response with digital emphasis 125 Hz to 16 kHz: within ± 0.1 dB (Track 12 Index 01 through Track 12 Index 05).

Linearity without dither 0 dB to -70 dB: within ±0.2 dB (Track 18 Index 01 to Track 18 Index 11).

Linearity with dither -80 dB to -90 dB: within ±1 dB (Track 19 Index 02 to Track 19 Index 03).

THD+N at 127, 997, and 10007 Hz: < 0.03 %
(Track 07 Index 02, Track 08 Index 01, and Track 09 Index 01)
(Note: 22-kHz band-limited measurement: < 0.02 %).

SMPTE intermodulation (IM) distortion: < 0.02 %
(Track 13 Index 01).

CCIF intermodulation (IM) distortion: < 0.01 %
(Track 13 Index 02).

Noise level re 1k-Hz reference voltage: < -90 dB
(Track 04 Index 01).

Crosstalk (non-driven channel re driven) 0.125, 1, and 10 kHz: < -100 dB
(Track 02 Index 01, 02, and 04 to Track 03 Index 01, 02, and 04)
(Note: measurements were made using the CROSSTALK function in the Audio Precision which band-limits its input signal from the non-driven channel to be centered about the frequency of the input signal from the driven channel.)

The operation of the DMR-4000 recorders was checked by playing back the 997 Hz harmonic distortion and the intermodulation distortion test tracks as recorded onto cassette. In all cases, the differences in the distortion results obtained using the CD directly and the cassette copy was at least an order of magnitude less than the results indicated above.

The DI-5 digital input board was replaced with the AD-23 analog-to-digital input board used to make the recordings. The digital outputs of the AD-23 board were connected directly to the digital inputs of the DA-15 board. The balanced analog outputs of the DA-15 board were connected directly to the 600-ohm inputs of the Audio Precision test system and the balanced analog 50-ohm outputs of the Audio Precision test system were connected directly to the analog inputs of the AD-23 board. The output voltages of the DA-15 board were nominally adjusted to 7.75 Vrms (+20 dB re $\sqrt{0.6}$ Vrms) and the AD-23 analog input stages were set for unity gain using the 1-kHz, 0-dB reference level on track 1 of the cassette copy of the CBS test disc, the measuring voltmeter in the Audio Precision, and a 1-kHz sine wave applied to the inputs of the AD-23 board. The results of these measurements were used to characterize the performance of the AD-23 analog-to-digital input board in combination with the DA-15 digital-to-analog output board as a function of frequency and/or amplitude. The frequency "sweeps" were actually performed at discrete frequency intervals (typically 100 points) over the frequency range 20 Hz to 20 kHz using the SWEEP TEST function in the Audio Precision test system. The amplitude "sweeps" were done at discrete amplitude points (typically 30 points) over a 20 decibel range using the SWEEP TEST function. Unless noted otherwise, the voltages at the analog inputs of the PCM-1630 during the measurements were nominally equal to the 1-kHz reference of 7.75 Vrms (+20 dB re $\sqrt{0.6}$ Vrms).

An overall summary of the results of these measurements follows:

Frequency response 20 Hz to 20 kHz: within +0.5/-1.0 dB

Phase response CH 2 re CH 1 20 Hz to 20 kHz: within ±3 deg

Linearity 0 dB to -70 dB re 1-kHz reference voltage: within ±0.4 dB

Linearity 0 dB to -90 dB re 1-kHz reference voltage: within ±3 dB

THD+N 20 Hz to 20 kHz at +14 dB re 0.775 Vrms (0 dBm): < 0.05 %

THD+N 20 Hz to 20 kHz at +4 dB re 0.775 Vrms (0 dBm): < 0.05 %

SMPTE IM distortion 0 dB to +14 dB re 0.775 Vrms (0 dBm): < 0.13 %

SMPTE IM distortion 0 dB to +12 dB re 0.775 Vrms (0 dBm): < 0.06 %

CCIF IM distortion 0 dB to +14 dB re 0.775 Vrms (0 dBm): < 0.04 %
(Note: the CCIF test signal was comprised of two test tones at frequencies of 11 kHz and 12 kHz combined in a 1:1 amplitude ratio.)

Noise level re 1-kHz reference voltage: < -90 dB

Crosstalk CH 1 re CH 2 (non-driven re driven) 20 Hz to 20 kHz: < -55 dB

Crosstalk CH 2 re CH 1 (non-driven re driven) 20 Hz to 20 kHz: < -70 dB
(Note: the crosstalk measurements were made using the CROSSTALK function in the Audio Precision which band-limits its input signal from the non-driven channel to be centered about the frequency of the input signal from the driven channel.)

The Fluke Model 8860 was used to establish the reference voltages present at the outputs and inputs of the PCM-1630 and set the gains of the analog input stages of the PCM-1630 to -1.0 dB during the recording of the direct and encoded copies of the selection master. The output reference voltages were set using the cassette-copy of the CD-1 test disc to playback the 1-kHz reference level recorded from Track 1 of the disc. The gains of the analog outputs of the PCM-1630 were adjusted to a level of +20.0 dB re $\sqrt{0.6}$ Vrms (+20.0 dBm) when terminated by the input impedances of their respective analog inputs in parallel with the nominal 10 megohm input impedance of the digital multimeter. When the Encoder was inserted between the inputs and outputs of the PCM-1630, the voltage levels present at the inputs to the Encoder was measured in the same fashion. In all cases, the variations in voltage levels present at the PCM-1630 outputs due, for example, to variations in load impedance or drift in the gains of the amplifiers in the PCM-1630 analog outputs were less than 0.05 dB during the direct or encoded recordings.

After the gains of the PCM-1630 outputs had been set, the noise at the inputs of the PCM-1630 was measured as a function of frequency using the Audio Precision test system. The noise at the inputs of the PCM-1630 was measured three times: once with the inputs connected directly to the output stages of the PCM-1630 and once with each of the two Encoders connected between the inputs and outputs of the PCM-1630. The equipment was connected and grounded as it was during recording of the direct and encoded copies of the selection master except that the test system inputs (in this case nominally 100 kΩ) were connected in parallel with the inputs of the PCM-1630. The noise was measured at 200 points over the range 20 Hz to 100 kHz using the SWEEP TEST function in the Audio Precision test system for each of the three cases, direct and two encoded. For frequencies in the range 20 Hz to 100 kHz, any differences in

the noise measured at the inputs of the PCM-1630 occurred more than 100 dB below the reference voltage of 7.75 V rms. For frequencies measured in the range 20 Hz to 4 kHz, the differences in the measured noise occurred more than 110 dB below the reference voltage of 7.75 V rms. Typically the differences in the noise were less than 6 dB at any given frequency over the frequency range measured.

Parallel Presentation

Copying, Encoding, and Editing Procedures

In preparing the material for presentation in the parallel listening study, the selections were copied digitally from compact disc to the selection master cassette using the same procedures as described above for the serial selections (see figure F-1). During the same time period that the selections were copied, a number of test signals were copied from the CBS CD-1 test disc for testing purposes. The recordings of the test material were then edited with the announcements of the start and end of the test selections using the PCM-1630 audio processor and DAE-1100A digital audio editor (see figure F-3). The announcements and test selections were edited such that there was a two-second pause between the announcements and the test material with the edit point(s) occurring in the middle of the two-second pause. During this editing process, the test signals recorded on the selection master were copied onto the same cassette at the conclusion of the parallel listening test. At no time was there any gain offset (as indicated by the illumination of the gain offset LED next to the GAIN OFFSET fader) introduced into the recordings. The parallel listening test signals and the test tracks were then copied digitally from the parallel master cassette onto two tracks of the PCM-3324 multi-channel recorder with the DMR-4000 player connected directly to the PCM-1630 audio processor (see figure F-4). (Note: the listening test signals and the test tracks were recorded onto two separate tapes using the INSERT record mode of the PCM-3324 with the connection of the recording equipment the same in each case.) The multi-channel tapes were "prestriped" (control and time code tracks prerecorded) using the DAE-1100A editor to generate time code with the word sync of the PCM-3324 "locked" to that of the PCM-1630 (SYNC CLOCK of the PCM-3324 set for external reference). All recordings were performed using a 44.1-kHz digital sampling frequency. All further recordings using the PCM-3324 were

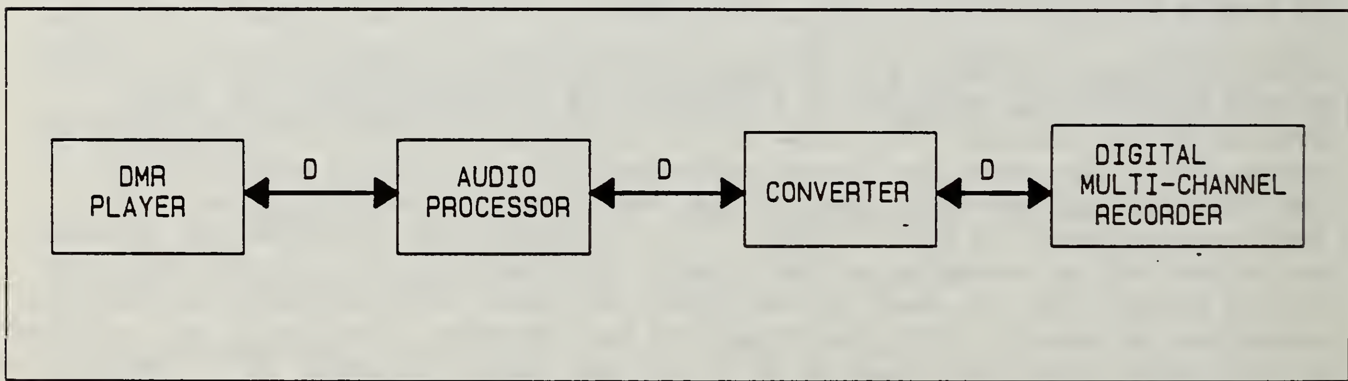


Figure F-4 Equipment setup for digital transfer to multi-channel digital audio tape. ("D" denotes digitized audio)

made at a 44.1-kHz sampling frequency with the PCM-3324 in the INSERT record mode and the SYNC CLOCK of the PCM-3324 referenced to that of the PCM-1630.

"Direct" and "encoded" recordings were made by connecting the analog outputs and inputs of the PCM-1630 in the same manner as described for the recording of the material used in the serial listening tests (see figure F-5). In this case, however, the entire test was recorded onto a set of tracks containing direct material and a set of tracks containing encoded material using the PCM-3324 to playback the initial digital recording of the listening test and record the direct and encoded versions of the listening test. The gains of the input stages of the AD-23 digital-to-analog board were checked and found to be set for -1.0 dB (see the previous section **Serial Presentation** regarding the reasons for this gain setting). Direct and encoded recordings were made of the test signals using the test-signal tape with the recording equipment connections the same as in the case of the recordings of the listening test except that only one Encoder was used in making these recordings. The test-signal recordings were made prior to the listening-test recordings and were used to verify that proper channel (left and right, or CH 1 and CH 2) and track assignments were maintained throughout the recording process. Encoder 012 was used to encode the first six selections in the test and Encoder 006 was used to encode the last four selections in the test. The AUTO PUNCH feature of the PCM-3324 was used to transition between the Prokofiev and Bernstein selections from material encoded using one Encoder to material encoded using the other Encoder. The transition point ("punch in" point, in PCM-3324 terminology) occurred during the pause between the announcement of the end of the Prokofiev selection and the announcement of the beginning of the Bernstein selection. To prevent the encoding of the announcements, the Encoder was switched IN during the two-second pause between the end of each of the announcements and the start of each of the selections and OUT during the two-second pause between the end of each of the selections and the start of each of the announcements The analog outputs on the PCM-3324 of the tracks containing the initial digital recording of the test were used to monitor the listening test when making the recordings.

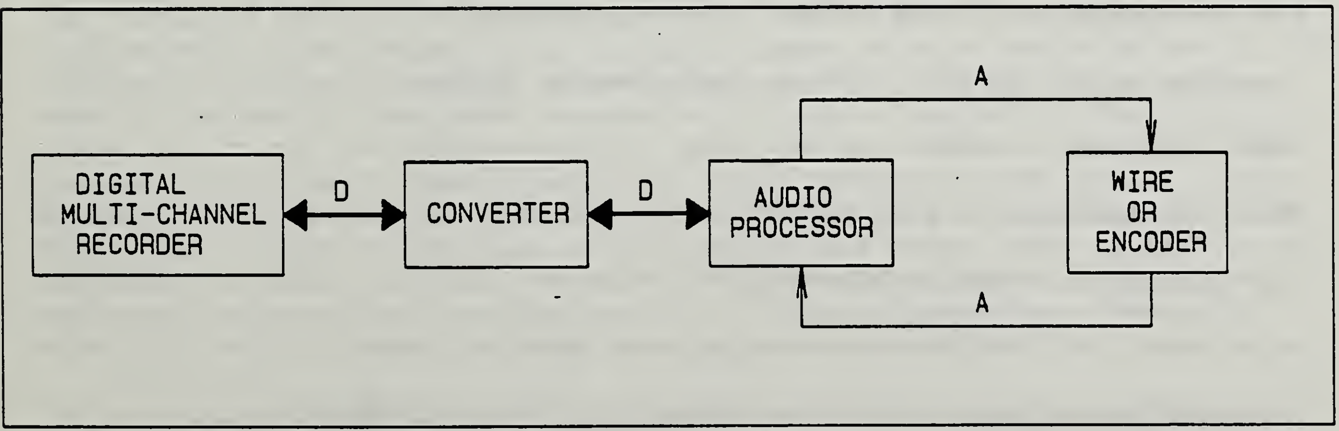


Figure F-5 Equipment setup for making direct and encoded tapes on the multi-channel recorder. ("D" denotes digitized audio; "A" denotes 2-channel analog audio)

The tracks of the direct and encoded material were then copied digitally onto three additional sets of tracks: a reference set (REF), A, and B. These recordings were made by directly connecting the balanced digital outputs and inputs of the tracks being used in the PCM-3324. First the tracks containing the direct material were copied on the REF set of tracks to insure that the reference material had been processed in the same manner as the material recorded on sets A and B. Then, since the majority of the direct recordings were assigned to A, all of the direct material was copied on the A set of tracks while, at the same time, all of the encoded material was copied onto the B set of tracks. In order to have the direct and encoded material randomly assigned to A and B, the track assignments of A and B were then reversed and three electronic edits were performed using the AUTO PUNCH feature of the PCM-3324. When performing these edits, the points at which the PCM-3324 began and stopped recording ("punch in" and "punch out" points, in PCM-3324 terminology) occurred during pauses between announcements of the beginning and end of selections. During these edits, the previously recorded material was replaced by the new material between the edit points. Prior to making these recordings, the test signals were recorded on the three sets of tracks REF, A, and B with the direct recordings assigned to REF and A and the encoded recordings assigned to B. The test signals were used to verify that proper track assignments were maintained throughout the recording process and to test the performance of the multi-channel recorder. The recorded material on tracks REF, A, and B was checked after the recordings were complete to verify that the encoding of the selections matched that of the test key and for the absense of switching transients during the times taken the Encoder was switched IN and OUT (the HP 3565S signal processing system and 330 computer mentioned in section 4.2 were used to perform this analysis).

Equipment Calibration

The same instrumentation and test procedures implemented when recording the material used in the serial study apply to recording the material used in the parallel study except with the additional testing of the PCM-3324 multi-channel recorder and the RS-422/TTL digital interface (converter). The digital interface was tested using a high-frequency function generator and oscilloscope prior to use. A number of measurements were made using the PCM-3324 to playback the digital copy of the test signals.

A summary of the results of these measurements follows:

THD+N at 997 HZ: < 0.01%

SMPTE intermodulation (IM) distortion: < 0.02 %

CCIF intermodulation (IM) distortion: < 0.01 %

Noise level re 1-kHz reference voltage: < -90 dB

Crosstalk (non-driven channel re driven) at 1 kHz: < -100 dB
(Note: measurements were made using the CROSSTALK function in the Audio Precision which band-limits the input signal from the non-driven channel centered about the frequency of the input signal from the driven channel.)

The Fluke Model 8860 was used to establish the reference voltages present at the outputs and inputs of the PCM-1630 and check the gains of the analog input stages of the PCM-1630 prior to making the direct and encoded recordings. The output reference voltages were set using the digital multi-channel copy of the 1-kHz reference level recorded from Track 1 of the CD-1 test disc. The gains of the analog outputs of the PCM-1630 were adjusted to a level of +20.0 dB re $\sqrt{0.6}$ Vrms (corresponding to +20.0 dBm) when terminated by the input impedances of their respective analog inputs in parallel with the nominal 10 megohm input impedance of the digital multimeter. When the Encoder was inserted between the inputs and outputs of the PCM-1630, the voltage levels present at the inputs and outputs of the Encoder was measured in the same fashion. In all cases, the variations in voltage levels present at the PCM-1630 outputs due, for example, to variations in load impedance or drift in the gains of the amplifiers in the PCM-1630 analog outputs were less than 0.05 dB during the direct and encoded recordings. The differences between the reference voltage levels measured at the inputs to the PCM-1630 prior to making the direct and encoded recordings with the Encoder switched IN or OUT were less than 0.2 dB.

Digital Error Generation, Correction, and Monitoring

The amount of digital error generation and correction during the playback and recording of the tapes used in the serial and parallel listening studies was visually monitored by watching the LED indicators on the DI-5 (digital input) board installed in the PCM-1630 used for recording, the CRC error and MUTE indicators on the PCM-1630 audio processors, the LED indicators on the DEC-15 (decoder) boards in the PCM-1630 audio processors, the CRC error indicators on the TBC (time base corrector) board in the PCM-3324, and the MUTE, HOLD, CON (concealment), COR (correction), and THR (through) indicators on the DCC (decoder control) board in the PCM-3324. Channel 1 (the factory default channel) was selected as the channel to be monitored by the indicators on the DCC board based on the assumption that, since it is one of the two digital channels nearest the edges of the tape, it would be most likely to indicate tape wear by increased error generation and correction (a digital signal of "0" was recorded on this channel when the tapes were prestriped using the REC MUTE function in the record mode selection section of the PCM-3324).

The status of the LED indicators on the DI-5 board during the copying of the test material from compact disc to digital audio cassette was with the DI SYNC indicator illuminated and the other indicators off. Two of the selections used in the serial study were recorded with digital emphasis. When playing these selections, the emphasis status indicator on the Denon Model DCD-3300 compact disc player became illuminated, which agreed with the status of the EMPHASIS indicator on the PCM-1630. There was never any illumination of the MUTE, ILLG (illegal), CRC (cyclic redundancy check code) error, M EMP (manual emphasis), or MONO indicators noticed during the time periods that digital audio was being copied (the function of the LED indicators was checked by observing their illumination when powering the PCM-1630).

The correction to the digital data that occurred during playback of the material recorded onto digital audio cassette was continuously monitored during the recording process through the use of the LED indicators on the DEC-15 board in the audio processor used to make the recordings. A decision was made

not to use the read-after-read function available through the use of the RAR-1 board; this decision was based on the possibility that the circuitry associated with the confidence head of the DMR-4000 playback machine might not be perfectly matched to the time delay in the circuitry associated with the record/play head and thus digital correction using the confidence head might introduce some audible effect with respect to time offset in corrected and uncorrected material when the test material was presented in pairs. Up to the recording of the S1 and S2 tapes associated with the serial tests, the digital errors that occurred (apparently on an intermittent and random basis) were totally corrected, as indicated by the DEC-15 board, with one exception when the digital audio data were averaged (interpolated) for an instant. During the production of the S1 and S2 tapes, an intermittent technical fault in the DMR-4000 recorder developed approximately 2.5 hours after starting the recording associated with each of the tapes S1 and S2. This fault may have been due to an intermittent problem with the SV-24 and/or CF-8 servo-control boards. When checking the S1 and S2 tapes during playback, each of them generated an excessive (continuous) amount of error correction in the DEC-15 board, starting approximately in the middle of each of the tapes (as indicated by the continuous illumination of the CRC error indicator on the audio processor). The error correction indicated by the status of the illumination of the LED indicators on the DEC-15 board in the audio processor was total correction (only the "Correction", or green, LED was illuminated) during the playback of the presentations of the test pairs. A representative of the manufacturer of the recorder and audio processor was contacted and confirmed that even though continuous error correction was occurring, if total error correction was indicated by the DEC-15 board then the errors occurring in the digital audio data should indeed be completely (totally) corrected. A third DMR-4000 mastering recorder was shipped immediately from Sony to NBS which, upon arriving at NBS, was used to make digitally corrected copies of tapes S1 and S2. The performance of this recorder was checked prior to use with the tracks recorded from the CBS test disc onto digital audio cassette used to measure THD+N at 997 Hz, SMPTE IM distortion, and signal-to-noise ratio. The administrator of the serial tests monitored the status of the LED indicators on the DEC-15 board installed in the PCM-1630 used during playback of the recordings prior to the completion of the corrected copies of S1 and S2. When making the digitally corrected copies of tapes S1 and S2, four instances of averaging, and four instances of averaging and/or holding were noted during the playback of the test pairs by the brief illumination of the "Average" (yellow) and/or "Hold" (red) LED indicators on the DEC-15 board (no occurrences of muting, as indicted by the illumination of the MUTE indicator on the PCM-1630, were noticed or heard). These eight points, along with the other point at which averaging of the digital data occurred during the production of tapes S1 and S2, were analyzed using the HP 3562A analyzer to compare the time records of the test pairs during which the averaging and/or holding was indicated with the time records of those test pairs in the same presentation in which total digital audio error correction occurred. A cursory analysis of low-pass filtered signals (6.25 kHz cutoff frequency) was performed by comparing the first time records at all nine of these locations in the digitally corrected serial tapes for both the left and right channels. (The low-pass filtering of the time-domain analysis was a result of the anti-aliasing filters of the signal analyzer being set such that the analog input signals to the analyzer were low-pass filtered at 6.25 kHz, effectively eliminating

frequencies above 6.25 kHz from the time record.) The results of this analysis indicated that the difference in the amplitudes of the first time record of the partially corrected audio data as compared with the first time record of the totally corrected audio data was less than 0.3 dB at any point in the time record for all nine points in the test tapes in which averaging and/or holding was noted.

The new (third) DMR-4000 was used to make the digital audio cassette recordings of the material used in the parallel listening study. The error correction during the recording and playback of the material for this study was monitored by observing the status of the indicators on the DEC-15 board in the PCM-1630, and the DCC and TBC boards in the PCM-3324. Only total correction was noticed during any occurrence of error generation, which occurred on an apparently intermittent and random basis. At the end and approximately the middle of the parallel listening tests, the sensitivity of the muting in the DCC board was set so that muting would occur when an error was continuous over 20 blocks of data (DIP SW1 switches 1 through 4 OFF -- "minimum sensitivity"). The tracks associated with REF, A, and B were listened to for the duration of the listening test and no occurrence of audible muting was noted (at the end of these listening sessions, the muting sensitivity setting on the DCC board was reset to the factory preset conditions, corresponding to muting occurring when an error is continuous over 50 blocks of data).

APPENDIX G

Audio Systems Used for Listening Tests

APPENDIX G: AUDIO SYSTEMS USED FOR LISTENING TESTS

The potential audibility of signal processing schemes intended for copy prevention is dependent upon the characteristics of the audio systems employed in the presentation of material encoded by such processing. It is not possible to address the question of audible effects upon the quality of the recorded material without using playback systems of high quality.

Since earphones and loudspeakers affect perceived sound quality in ways that are often quite different, and since some listeners may express a strong preference for one or the other in critical listening sessions, audio systems using earphones as well as loudspeakers were available for the serial listening tests. Earphones were used for the parallel listening tests.

From the measurements that had been performed upon the Encoders and preliminary listening studies by NBS staff members, it was apparent that the principal, potentially audible effects upon material passing through the Encoder would be attributable to the gain and phase characteristics of the encoding filter at frequencies near 3840 Hz. It was considered essential that the frequency response characteristics (in both amplitude and phase) of the audio playback systems be relatively smooth functions of frequency within and near this range. Sharp, large peaks and/or dips at these frequencies might exaggerate or obscure the potentially audible effects of the encoding filter. Distortion, particularly intermodulation distortion, also needed to be acceptably representative of high quality audio systems.

Playback Systems

Functional diagrams of the playback systems used during the serial and parallel presentations of recorded material are shown in figures G-1 and G-2. The compact disc player, synthesizer, and recording equipment used to produce the recordings for the listening sessions are described in Appendix F. A typical two-channel-headphone station (two-channel amplifier/polarizing voltage supply plus headset comprising earphones and the means to hold them on the head) and the amplifier/loudspeaker system used during the tests are described below. The DMR player (digital master recorder), multi-channel recorder, converter (digital interface), and the audio processors are the same type used in making the recordings.

In the case of the serial playback system, the composite digital outputs and input of the DMR player were connected to the audio processor, and the analog outputs of the audio processor were connected in an unbalanced configuration to a two-channel attenuator in parallel with the inputs of the two-channel power amplifier.

For the parallel presentation, the balanced digital outputs of the multi-channel recorder were connected to a custom (not mass produced) switching network to allow switching between pairs of tracks on the multi-channel recorder. The balanced digital outputs of the switching network were changed to unbalanced digital signals by the converter, and the unbalanced digital outputs of the converter were connected to the audio processor. The analog outputs of the audio processor were connected in an unbalanced configuration

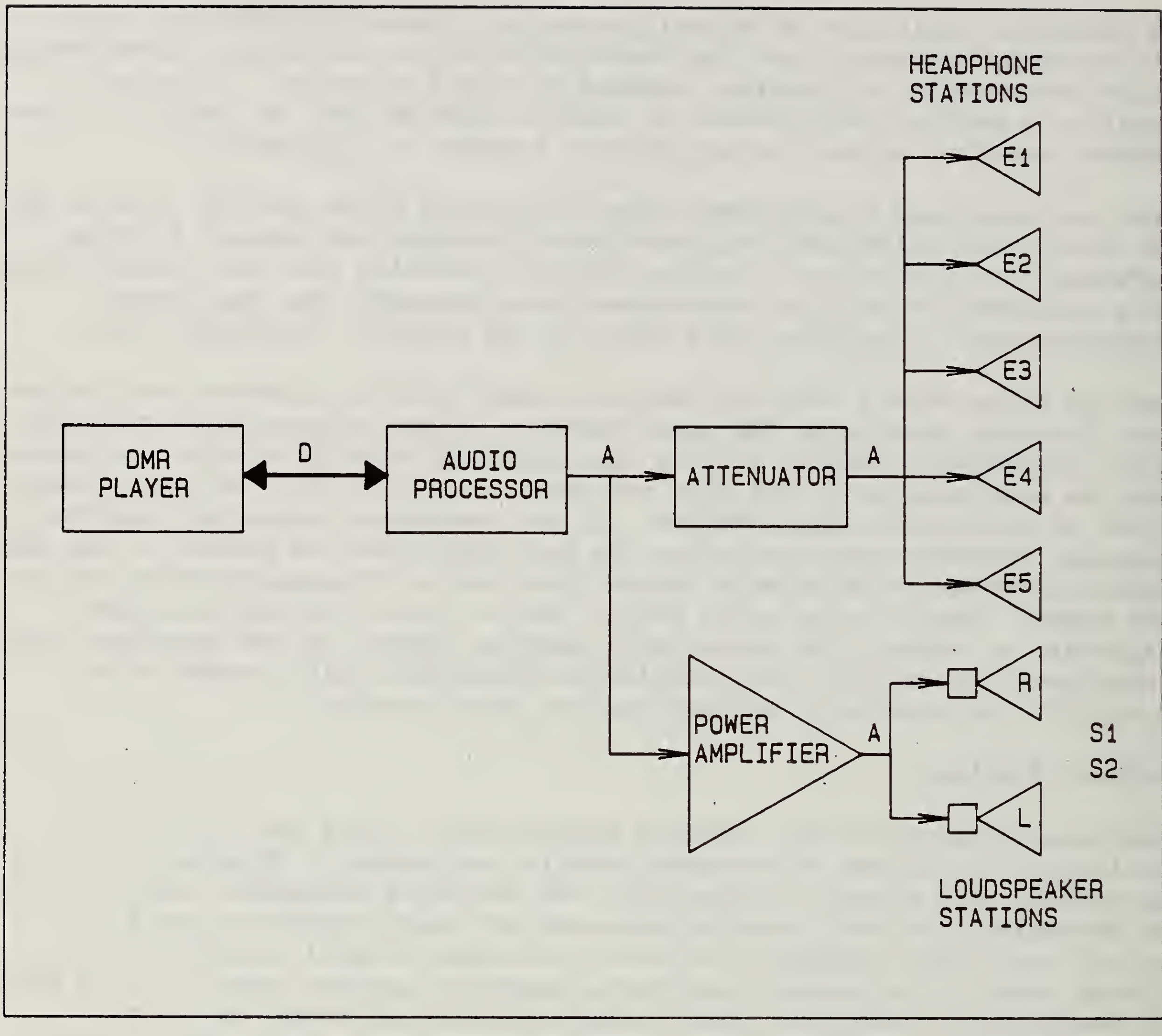


Figure G-1 Equipment setup used for audio playback in serial listening tests. ("D" denotes digitized audio signal. "A" denotes 2-channel analog audio signal.)

to a two-channel stepped attenuator which allowed the listener to adjust the volume of the test material during playback. A remote control was provided to allow the listener to manipulate the tape transport, which was located in a room adjacent to the room used for listening. The meter control (MT-16) board was removed from the audio processor during the parallel testing in order to eliminate the action of the peak indicating meters in the audio processor. The recorder, converter, and audio processor used to present the parallel test material were the same units used to produce the recordings for the test. Note that both the serial and parallel tapes were made and reproduced in such a way that at no time was the listener comparing differences introduced into the test material by the direct and encoded material having been processed by two different sets of analog-to-digital or digital-to-analog converters, or analog amplifiers (except those in the Encoder).

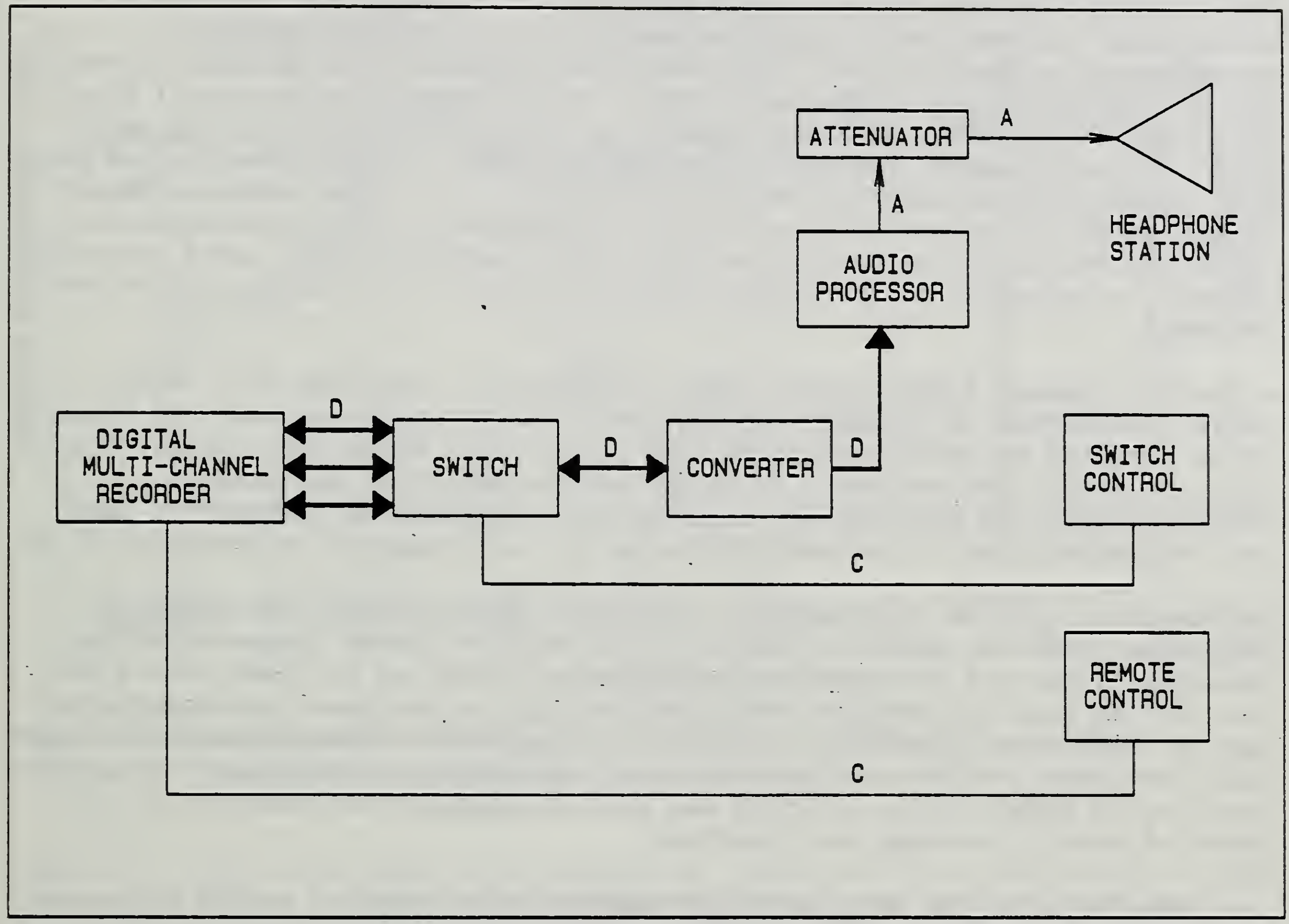


Figure G-2 Equipment setup used for audio playback in parallel listening tests. ("D" denotes digitized audio signal. "A" denotes 2-channel analog audio signal. "C" denotes control signal.)

Earphones

There are no established national or international standards for measuring the performance of earphones intended for reproduction of high quality stereophonic recordings. The measurements reported here were performed by placing the earphones upon an anthropometric manikin known as KEMAR (an acronym for the Knowles Electronic Manikin for Acoustic Research) [1,2].[1] It is fitted with a modified four-branch Zwislocki coupler [3] containing a microphone (B&K Type 4134) that measures a sound pressure approximating that at the eardrum of

[1] Figures in square brackets indicate the literature references at the end of this Appendix.

a typical adult as well as the current state of the field permits. The coupler used for these measurements was placed in the right side of KEMAR for measurements with the right earphones of the five headsets used in this study, and was placed in the left side for the measurements with the left earphones. This procedure prevented any slight differences due to normal tolerance differences between couplers from appearing as a difference between measurements performed on the left and right earphones of a headset. Each of the five headsets was of the same type: Stax SR-Lambda (Professional) Electrostatic Earspeakers, with the polarizing voltage and audio driving amplifiers for each headset provided by a STAX SRM-1/MK-2 Professional Driver Unit. Each headset/driver unit was tested as a system, i.e., the "earphone" measurements that are discussed here relate the inputs to the driver unit (which in the listening systems are the analog outputs of the audio processor, suitably attenuated) to the output sound pressures at the coupler microphone in KEMAR.

A Hewlett-Packard 3562A dynamic signal analyzer [4] was used with "burst chirp" excitation (a sinusoid rapidly swept in frequency) so that the total chirp duration was well within the time record upon which Fast Fourier Transform processing was performed to determine the amplitude and phase characteristics of each set of earphones as a function of frequency. Each such characteristic is defined by data points at frequency intervals of 25 Hz.

Varying the position of a headset upon KEMAR indicated that the smoothest amplitude-frequency characteristics at and near the notch frequency of the encoding filter are obtained with the headset pushed as far back toward the rear of the head as possible and placed as high on the head (although height was not nearly as significant in effect) as possible. The characteristics of each headphone station were measured and each station was balanced using this position on KEMAR. This position was also recommended to listeners at the start of their listening test sessions.

For each headset, the left and right earphone sensitivities at 800 Hz (a frequency at which relatively replicable measurements could be performed upon KEMAR) were matched within 0.7 dB. The sensitivities of all the earphones of four of the headsets were within 0.6 dB of each other. Including the fifth headset (used only when necessary), the sensitivities of all the earphones of all headsets were within 2.4 dB.

Figure G-3 shows the amplitude-frequency characteristics of the left and right earphones of a typical headphone station. The resonances just above 100 Hz and the decrease in amplitude as frequency decreases below 100 Hz are considered to be influenced by an artifact due to the use of KEMAR with earphones of this type. At least some of the differences between earphones at frequencies above perhaps 4 kHz are attributable to an intentional slight asymmetry between the left and right ears of KEMAR. Measurements performed upon KEMAR at these higher frequencies may not be representative of the performance of the earphones upon an individual adult because of differences between the manikin and different individuals. The significant aspect of these measurements is the smoothness of the amplitude-frequency characteristics and the excellent match between left and right earphones near 3840 Hz.

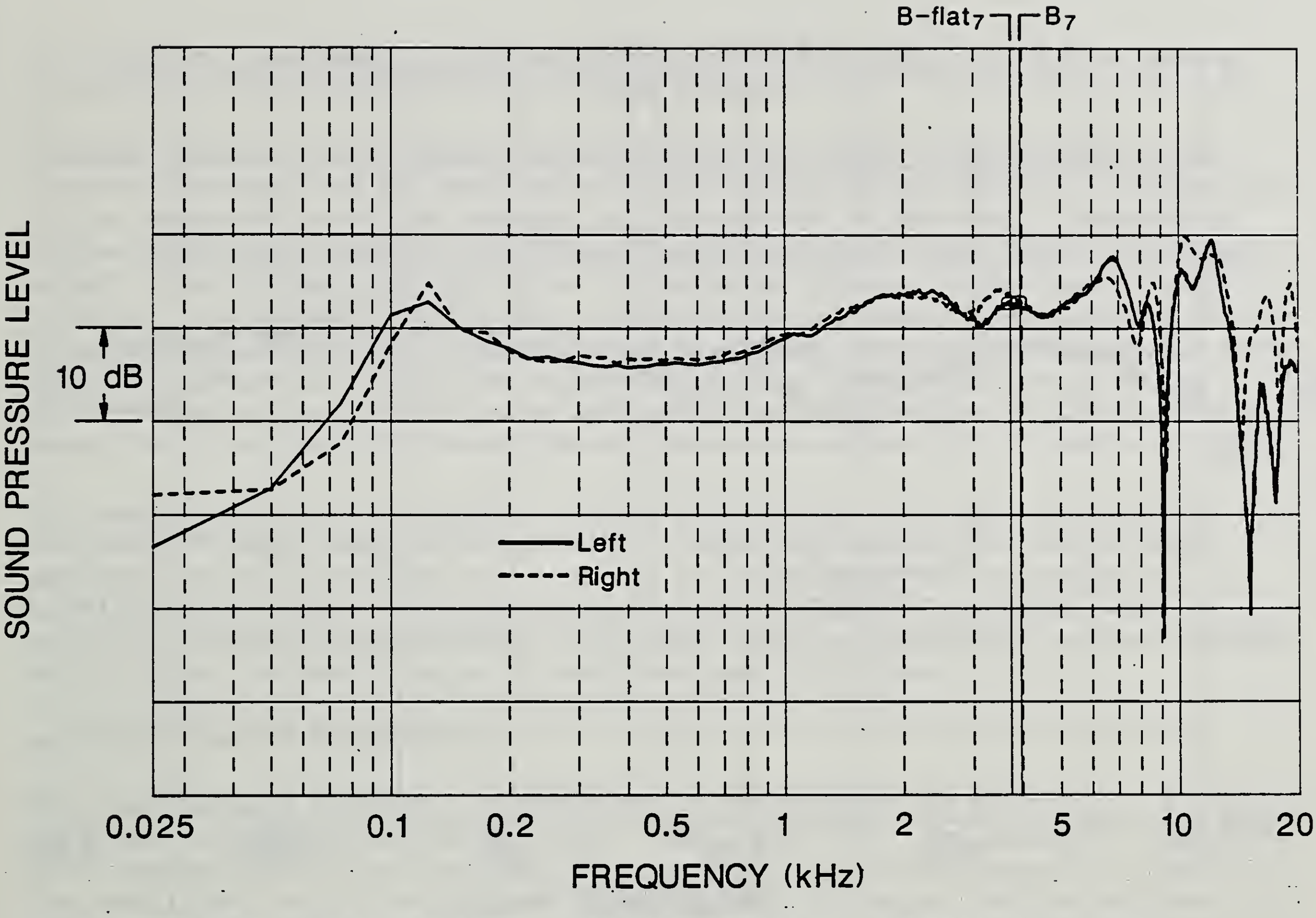


Figure G-3 Amplitude-frequency characteristics of the left and right earphones of a typical headset.

Figure G-4 shows the phase characteristics (relative to an arbitrarily chosen reference) of the left and right earphones of the same headset, with the frequencies displayed limited to the range for which artifacts of measurement are not likely to be a potentially dominant factor. Most significant are the smoothness of the characteristics and the excellent match between earphones at frequencies near the encoding filter notch frequency.

Intermodulation distortion was measured using signals specified by the Electronic Industries Association of America (EIA), which are available on Track 13 of the CBS CD-1 Test Disc. This track contains signals for the "Twin-Tone (CCIF)" method, consisting of simultaneously presented sine-wave signals of equal amplitude at 11 kHz and 12 kHz, and signals for the "SMPTE" method, consisting of simultaneously presented sine-wave signals at 60 Hz and 7 kHz with amplitude ratio 4:1, respectively. The signals from this track were obtained at the analog outputs of a compact disc player (Sony CDP-610ES) that was known from prior measurements at these outputs to produce a very low-distortion output from this track of the CD-1 Test Disc. The dynamic signal analyzer (HP 3562A) was used with its source turned off and with the "flat-top" window suitable for the accurate determination of the amplitudes of sine-wave signals (including distortion components) in the measured amplitude-frequency characteristics.

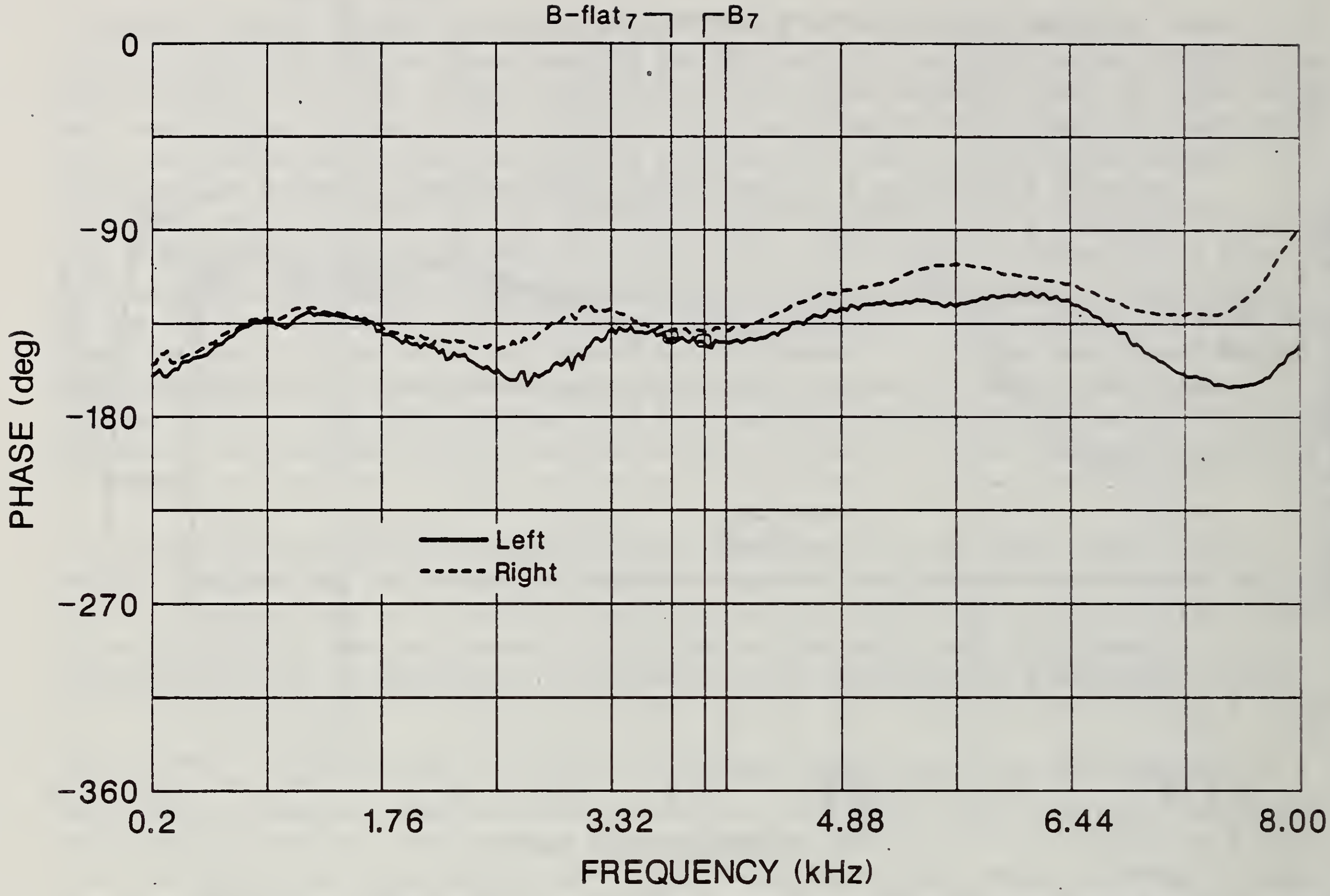


Figure G-4 Phase characteristics of the left and right earphones of a typical headset, relative to an arbitrary reference.

For the "Twin-Tone (CCIF)" signal, each distortion component (including the difference component at 1 kHz) produced by each earphone of the typical headset was at least 60 dB lower in SPL than the sine-wave signals presented at 11 kHz and 12 kHz. Considering that each of these signals resulted in a very high SPL (at least 104), the measured performance of the earphone system was very good.

For the "SMPTE" signal, each distortion component was typically at least 58 dB lower in SPL than the sine-wave signal presented at 7 kHz. Considering that the 7-kHz signal resulted in a high SPL (e.g., 93), and that the sine-wave signal at 60 Hz was presented to the input of the earphone system at a level 12 dB greater than the level of the 7-kHz signal, the measured earphone system performance was again very good.

Loudspeakers and Listening Room

A decision was made not to use multiple-driver loudspeaker systems having a crossover between drivers at a frequency between 3 kHz and 5 kHz, to avoid any confusion between possibly-audible effects of the encoding notch filter and possibly-audible effects of the speakers' crossover networks. The Energy 22

Pro Monitor loudspeakers that were used are two-way systems with a single crossover point, near 1500 Hz, between drivers.

Several types of measurements were performed on these loudspeakers in an anechoic chamber and in the room used for the listening studies. Measurements in the large NBS anechoic chamber were performed to determine loudspeaker system amplitude-frequency and phase characteristics (both axial and non-axial), distortion, and the degree of matching between loudspeakers under precisely controlled conditions. Measurements of room properties and of loudspeaker impulse response and amplitude-frequency characteristics in the room used for listening studies provided measures of loudspeaker performance in the listening room, as influenced by the acoustical properties of the listening room.

All measurements of loudspeaker system performance reported here include the two-channel amplifier (McIntosh MC7270) that was used to drive the loudspeakers in the listening tests, i.e., the amplifier and loudspeakers were considered as a unit to comprise the device for which properties were being measured. Previous measurements indicated that this amplifier was performing well within its manufacturer's published specifications.

Anechoic Chamber Measurements

For the anechoic chamber measurements, a loudspeaker was placed 2.7 m from the reference microphone Bruel and Kjaer (B&K) Type 4133, followed by a B&K Type 2645 preamplifier, and B&K Type 2607 and 2606 laboratory measuring amplifiers. The amplified output voltage of the reference microphone system was measured with a dynamic signal analyzer (HP 3562A), with the insert-voltage technique used to measure and to correct for the slight departures from uniform amplitude-frequency and phase characteristics of the preamplifier-measuring amplifier combination.

Initial measurements of each loudspeaker system's amplitude-frequency and phase characteristics and the impulse responses obtained from these characteristics by inverse FFT calculation, both with and without the loudspeaker protective grille covers in place, showed that removing this cover resulted in smoother characteristics in the frequency range of interest and a better impulse response for each loudspeaker. NBS staff members also performed subjective listening assessments of the loudspeaker systems in the listening room with music reproduced from compact discs, and concluded that the systems produced a slightly higher perceived quality of reproduction with these grille covers removed. The grille covers were removed for all listening tests and for all measurements reported in this study.

Figures G-5 and G-6 show the amplitude-frequency and phase characteristics, respectively, measured on-axis, for the loudspeakers in the anechoic chamber, using a "burst chirp" source and FFT processing in the dynamic signal analyzer. The lowest frequency shown for these measurements is 100 Hz, principally because the influence of typical listening rooms upon loudspeaker response is so dominant at frequencies below a few hundred hertz that there was little point to performing measurements under anechoic conditions at very low frequencies. The amplitude-frequency characteristics for the left and

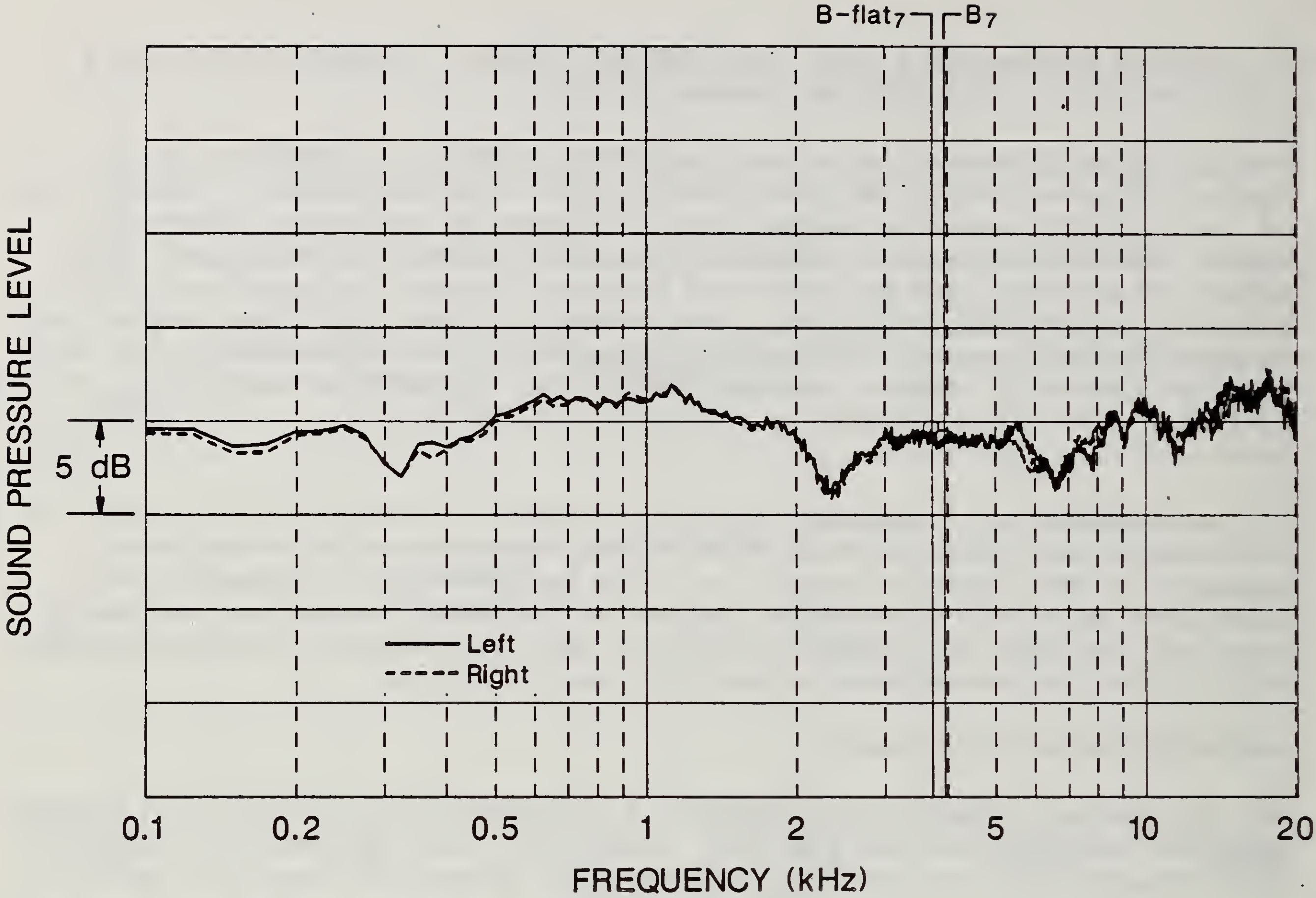


Figure G-5 Amplitude-frequency characteristics of the two loudspeakers, as measured in the large NBS anechoic chamber.

right loudspeakers are within a few tenths of a decibel of each other at most frequencies, and both are flat within +4/-3 dB at all frequencies of measurement, and flat within ±1 dB at frequencies from 2.7 kHz to 6.0 kHz, the range that includes the encoding filter notch frequency. The phase characteristics, plotted on a linear frequency scale so that a pure time delay would correspond to a straight line, are smooth functions of frequency, and the characteristics for the left and right loudspeakers are almost identical, any differences probably being primarily attributable to the delay time associated with the uncertainty (a few mm) in the positioning of the loudspeakers in the anechoic chamber.

The amplitude-frequency and phase characteristics were also measured in the anechoic chamber for each loudspeaker at other angles than axial. Over the range ±45 deg relative to axial, these characteristics remained relatively smooth functions of frequency, particularly for frequencies at and near the notch frequency of the encoding filter.

Intermodulation distortion was measured for each loudspeaker in the anechoic chamber using the same signal source and dynamic signal analyzer used for the earphone measurements described previously. For the "Twin-Tone (CCIF)" signal, each distortion component was found to be at least 56 dB lower than

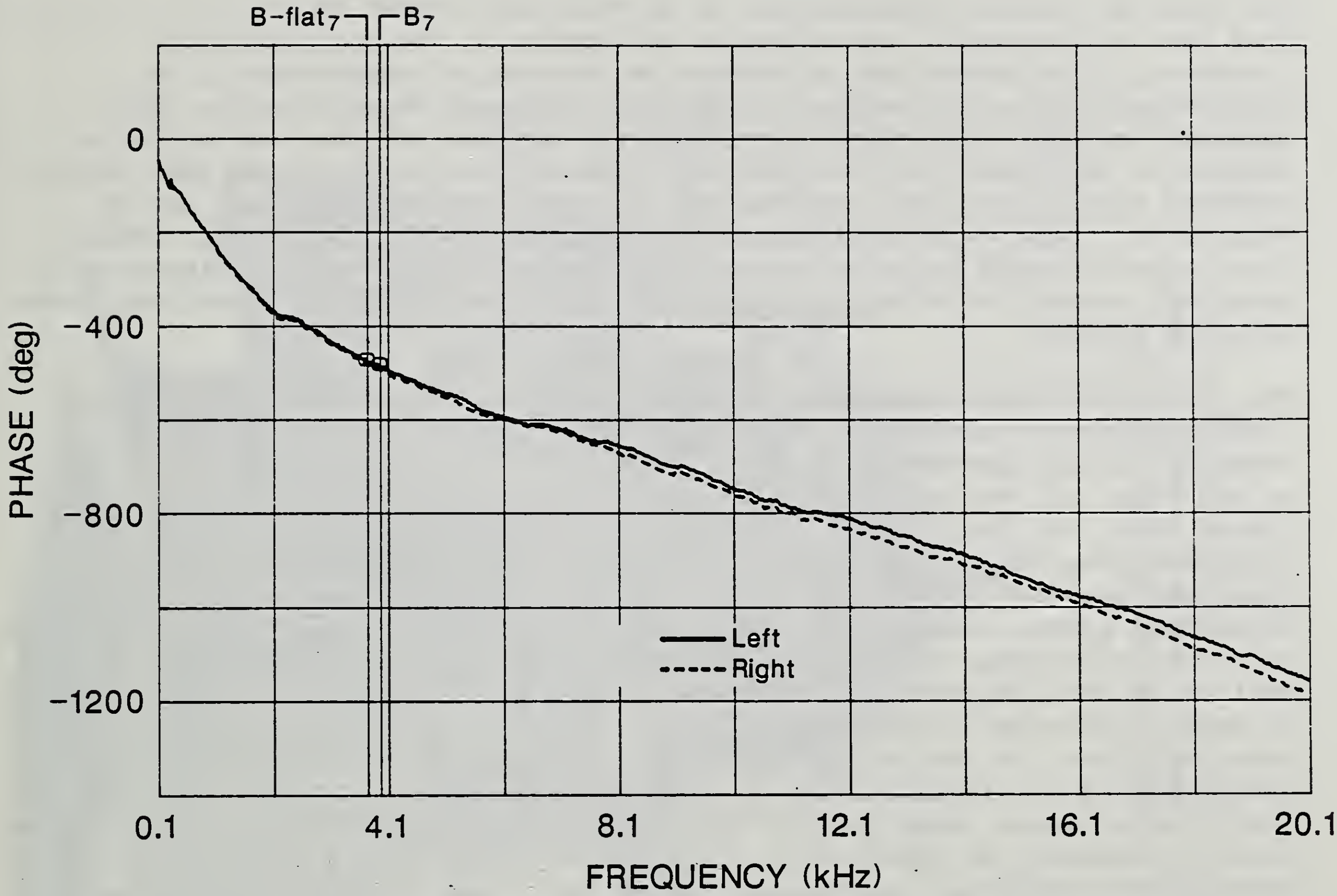


Figure G-6 Phase characteristics of the two loudspeakers, as measured in the large NBS anechoic chamber, relative to an arbitrary reference.

the level of the closer of the tones at 11 kHz and 12 kHz. For the "SMPTE" signal, with the overall sound pressure level being 80 dB (about 89 dB at one meter, a distance often used in such measurements), each distortion component was at least 40 dB lower than the level of the sine-wave signal presented at 7 kHz.

Listening Room Measurements

The listening room had a rectangular floor plan, 4.9 m x 6.7 m, with a drop ceiling 2.44 m from the floor. The floor was plush carpet and carpet pad over a concrete slab on grade. The walls were concrete block with a hard, smooth finish. Two windows (double glazed) on one long dimension of the room were covered with rather lightweight curtains. The loudspeakers were placed upon solid concrete blocks near one short wall of the room, and a rubber pad was placed beneath the front end of each loudspeaker, so that a normal to the front of the loudspeaker at the center of the high-frequency driver formed an angle of approximately 5 degrees to the floor. These normals then passed at the approximate ear level of two listeners seated side by side with their ear positions about 2..7 m from the line joining the front surfaces of the loudspeakers. The listeners were centered between the two loudspeakers, so that each was as nearly equidistant from the speakers as possible.

The room was treated acoustically to be relatively sound absorptive in the end near the loudspeakers, and diffusely reflective in the end behind the listeners. The intent was to provide an interval of approximately 10 milliseconds after the arrival of directly radiated sound from the loudspeakers in which reflections from interior surfaces of the room would be minimal in amplitude, so that the localization cues contained in this directly radiated sound during the reproduction of stereo recordings would not be obscured by these early reflections from room surfaces. Subsequent reflections were intended to be relatively diffuse, so that glaring reflections of transient sounds and severe standing wave patterns excited by sustained notes could be avoided.

The acoustical treatment of the room included the use of diffusing panels ("QRD Diffusors" manufactured by RPG Diffusor Systems, Inc.) and broad-bandwidth sound absorptive units ("Abffusors" from the same source), as well as one-inch and two-inch thick fiberglass panels. Diffusing panels were placed along the short walls of the room, and the "Abffusors" were placed along the long walls from the diffusors behind the loudspeakers, to the approximate position of the listeners. Either "Abffusors" or two-inch thick fiberglass panels were mounted as ceiling tiles in positions chosen to prevent specular reflections of directly radiated sound from the loudspeakers off the ceiling to the listeners. A fiberglass panel was placed on the floor in front of each loudspeaker, with the end farthest from the loudspeaker elevated by a concrete block, to reduce specular reflections from the floor. The diffusing panels behind the loudspeakers, as well as most of the remaining wall surface above these panels, were covered with fiberglass panels to further improve the perceived quality of sound reproduction, and, in particular, the ability to distinguish individual instruments and/or individual notes in rapidly articulated musical passages. Additional fiberglass panels were placed just above the loudspeakers and within the room as needed to attempt to eliminate flutter echoes between parallel surfaces, annoying standing wave patterns, and potentially deleterious specular reflections. The final configuration of the room is shown in figures G-7 and G-8.

The efficacy of the acoustical treatment was checked by using a calibrated microphone system (B&K Type 4133 microphone, Type 2645 preamplifier, and Type 2607 measuring amplifier) and an audio spectrum analyzer (Techron Division of Crown International, Model TEF-10) to measure the temporal pattern of directly radiated and reflected sound, with the microphone at the listening position. These and other measurements verified that the intended results had been achieved.

Reverberation times were measured in octave bands using "pink noise" produced by a sound power source (B&K Type 4205) placed between the two loudspeaker systems so that it was equally distant from each, but 15 cm farther from the listening position than a line joining the front faces of the loudspeakers. The receiving microphone was placed at the listening position, at a point between the positions of the listeners, but in the absence of the listeners. The receiving microphone system (B&K Type 4166 microphone, Type 2639 preamplifier, Type 2607 measuring amplifier, and Type 1617 bandpass filter) output was recorded on a level recorder (B&K Type 2305) to produce paper strip-chart type records from which the reverberation times for various octave

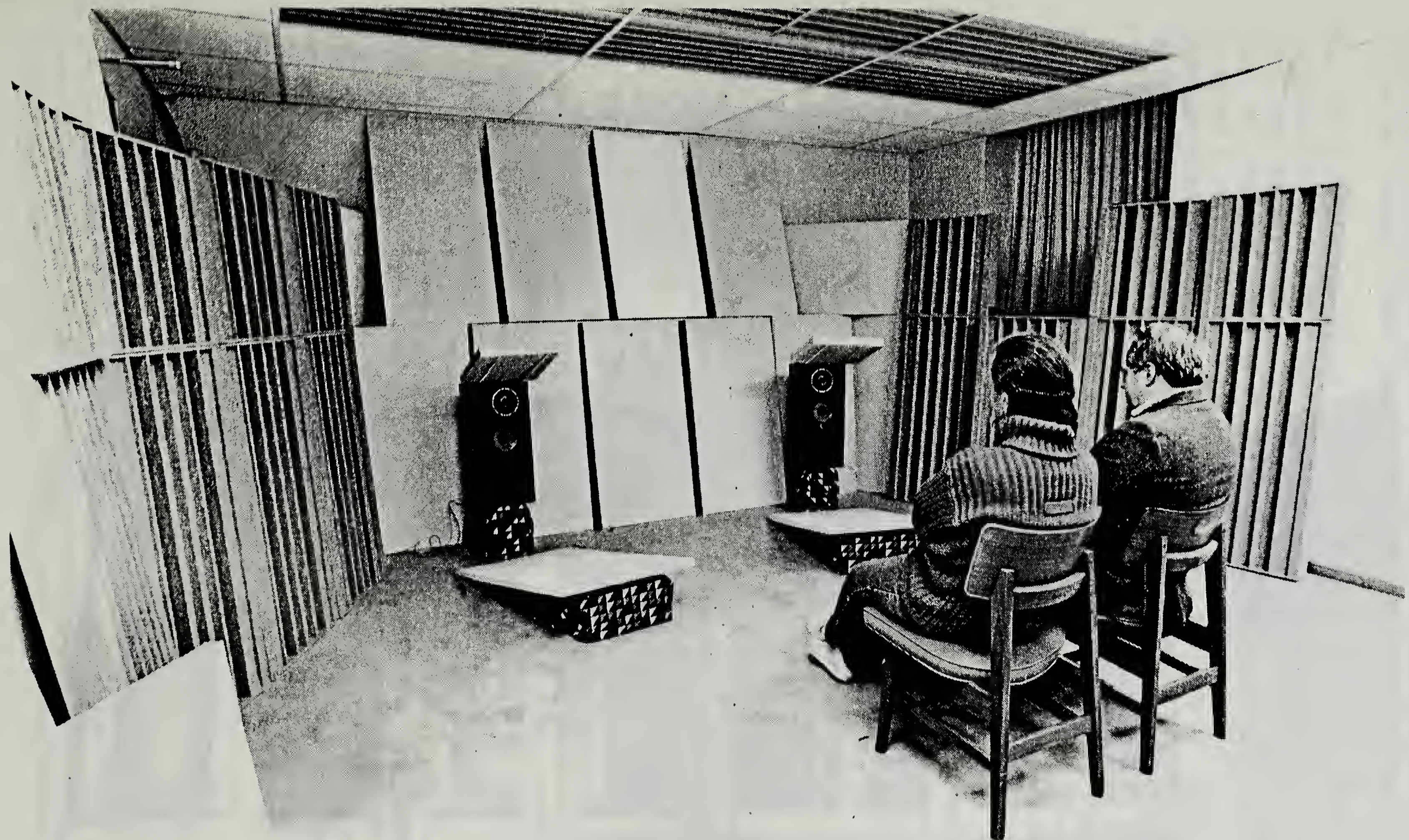


Figure G-7 The room used for the serial listening tests. The two subjects are shown facing the loudspeakers and the absorptive material used to minimize reflections from the diffusors and wall behind the speakers and from the side walls.

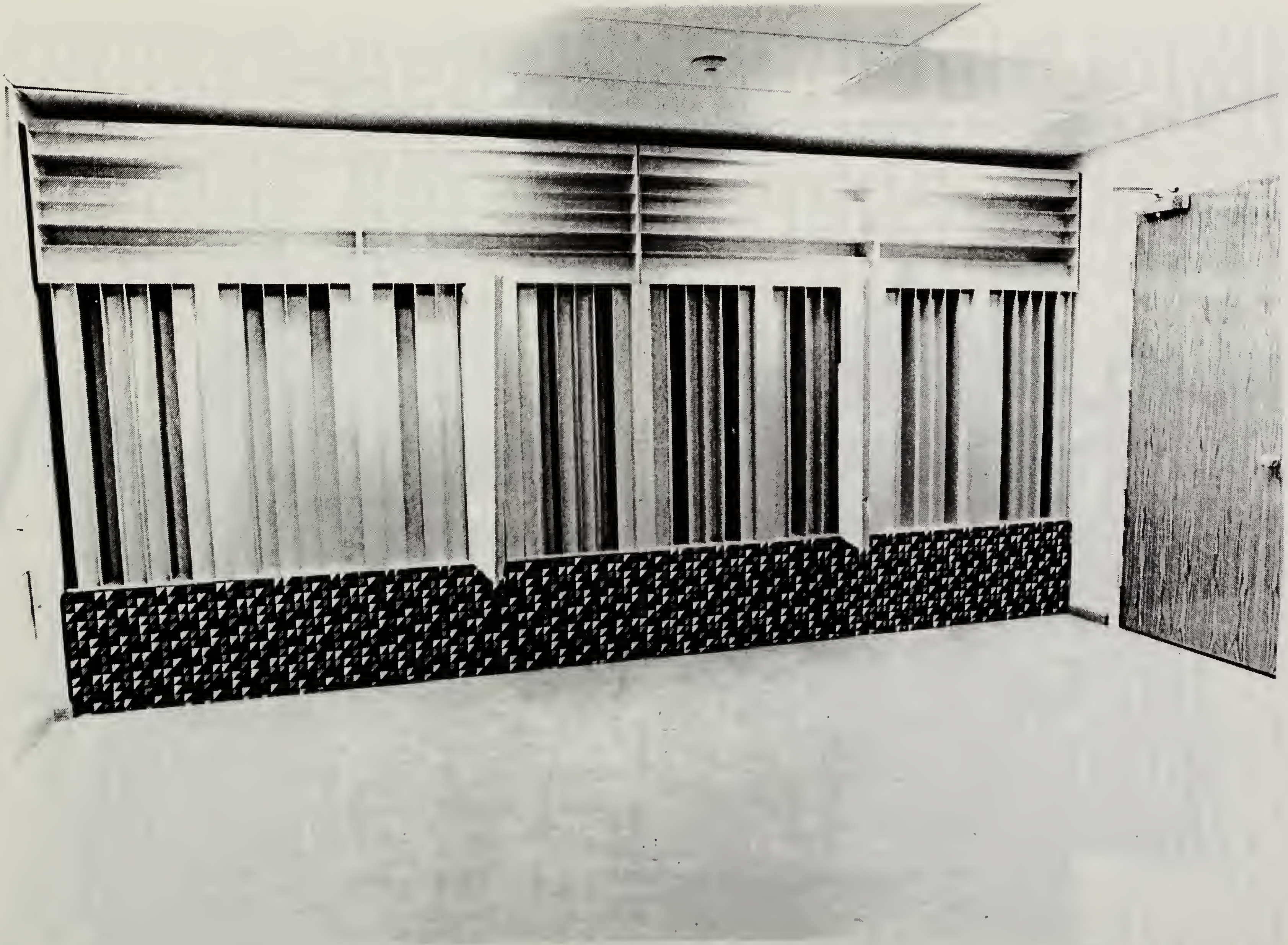


Figure G-8 Diffusing panels at the back of the room used for the serial listening tests.

bands could be determined. For the octave band centered at 125 Hz, the reverberation time was about 0.3 second. For each of the octave bands centered at 250 Hz through 8 kHz, the reverberation time was about 0.2 second.

The critical distance (the distance, from the sound source, at which the sound pressure level due to directly radiated sound is equal to the sound pressure level due to reflected sound) was determined to be about 1.8 m at frequencies near the center frequency of the encoding notch filter. Thus, the listeners received a substantial contribution of directly radiated sound, which is considered helpful for careful monitoring sessions in which it is intended to perform critical judgments that are not excessively influenced by the particular acoustical characteristics of the listening room.

An additional measurement of loudspeaker/room behavior was performed on each loudspeaker sequentially by using a dynamic signal analyzer (HP 3561A) to generate an audio-frequency-band-limited impulse [5], which was connected to the loudspeaker system. The signal analyzer was then triggered with the appropriate delay parameter setting to capture the first 10 msec of the impulse response of the loudspeaker (measured at the listening position using the same microphone system that had been used for the measurements in the anechoic chamber) as a low-pass-filtered (for anti-aliasing purposes), digitized time record that could be displayed and plotted as an analog waveform. This waveform (with uniform window weighting) was subjected to FFT processing to obtain the amplitude-frequency characteristic of the loudspeaker. Since this characteristic includes not only the direct sound radiation from the loudspeaker, but also the reflections from room interior surfaces occurring within the first 10 msec of the impulse, it is a measure of the loudspeaker/room-reflection amplitude-frequency characteristic that should be less smooth than, but essentially similar to, the corresponding anechoic response at frequencies above about 1 kHz, where the acoustical treatment of the room is considered particularly effective.

Figure G-9 shows the time record waveforms measured as described above for the left and right loudspeakers in response to the source waveform (which also exhibited a dominant initial negative peak). Reflections are relatively small in amplitude and are essentially negligible by the end of the time record.

Figure G-10 shows the measured amplitude-frequency characteristics obtained by FFT analysis of the time record waveforms for the left and right loudspeakers in the listening room. The characteristics are well-matched, flat to within ±4 dB from 2.9 kHz to 5.9 kHz, and are even better at frequencies near the encoding notch frequency. The characteristics are, as expected, somewhat roughened by room effects, but are comparable to the amplitude-frequency characteristics (figure G-5) measured in the anechoic chamber at corresponding frequencies.

A measurement of the loudspeaker/room behavior that includes the effects of all room reflections in response to a continuously applied broad-band signal was also performed. Each loudspeaker was excited sequentially (i.e., first one loudspeaker alone, then the other alone) by the "pink noise" output of a generator (GenRad 1382) of (statistically) stationary noise, and the loudspeaker amplitude-frequency characteristic was measured in

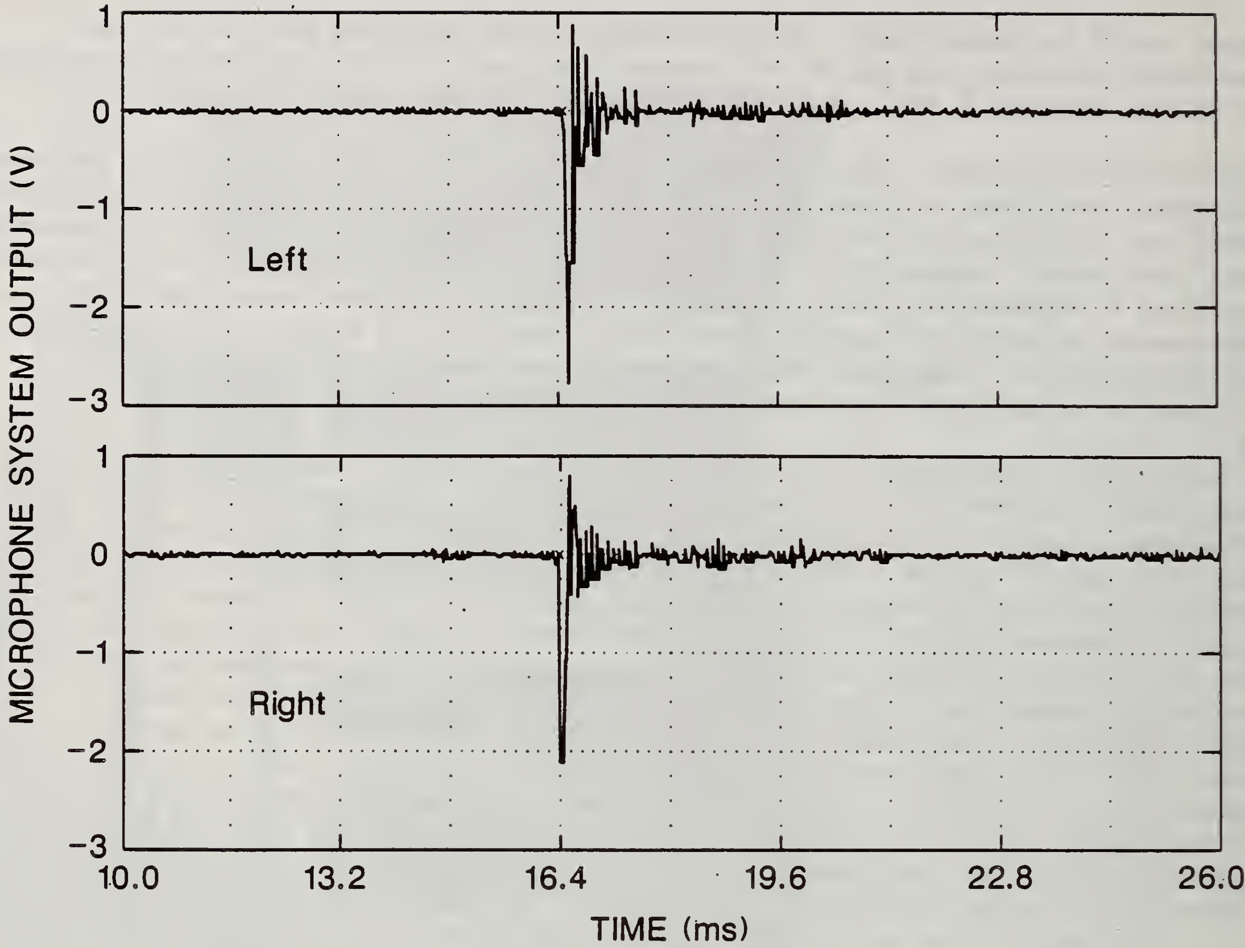


Figure G-9 Impulse response of the two loudspeakers as measured in the listening room.

one-third-octave bands using the analyzer in its synthesized one-third-octave-bands mode [5], with the microphone system (B&K Type 4134 microphone, Type 2639 preamplifier, and Type 2607 measuring amplifier) determining the sound pressure level at the listening position. Ideally, these measurements would result in essentially "flat" (uniform with frequency) amplitude-frequency characteristics (but note that controlled departures from a flat characteristic are also often advocated, e.g., a characteristic that gradually decreases in amplitude with increasing frequency to simulate the sound received at many good seats in typical concert halls). With the same level of noise presented sequentially to each loudspeaker, these characteristics for the left and right loudspeakers at the one-third-octave band centered at 4 kHz (containing the notch frequency of the encoding filter) were equal to the nearest 0.1 dB. Both characteristics were also within +3/-4 dB of their value at 4 kHz for all one-third-octave bands centered at frequencies from 315 Hz to 12.5 kHz, and within +7/-8.5 dB of their values at 4 kHz for all bands from 30 Hz to 20 kHz.

The spatial uniformity of the sound field produced sequentially by each loudspeaker within the listening region (range of possible ear positions of

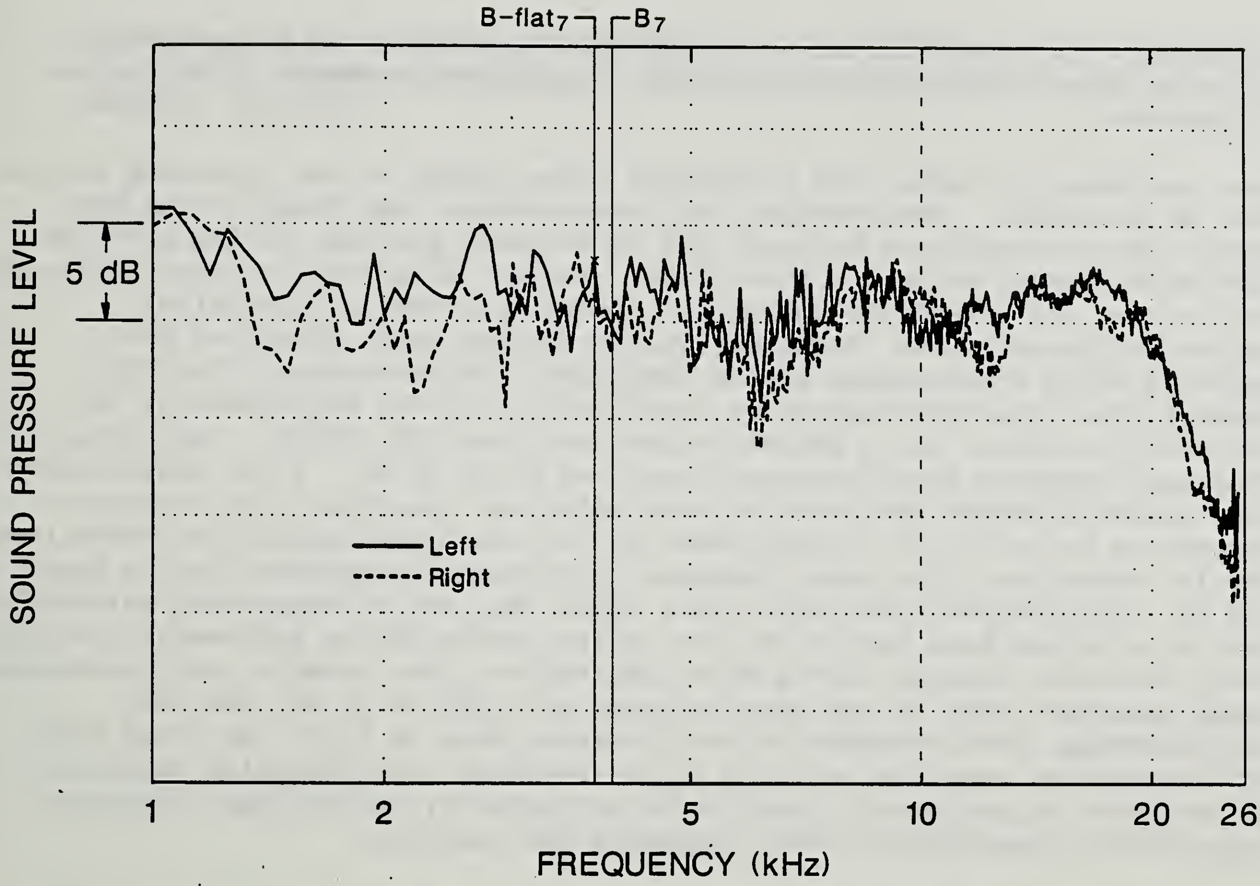


Figure G-10 Amplitude-frequency characteristic of the two loudspeakers as determined in the listening room by FFT analysis of time record waveforms.

the listeners in the absence of the listeners) in response to "pink noise" was measured with a sound level meter and octave-band filter (B&K Type 2204 with Type 4166 microphone and Type 1613 Octave Filter Set). Using the C-frequency-weighting network (but no additional filtering) and the "slow" exponential averaging time on this instrument, the difference between the highest and lowest sound pressure levels measured throughout the listening region was 0.8 dB and 1 dB for the left and right loudspeakers, respectively. Using the octave-band-filter (but no other frequency-weighting network) and "slow" exponential averaging time, these differences, for either the left or right loudspeaker, at each indicated center frequency of the octave band filter, were: 0.7 dB for the 1-kHz octave band and 0.5 dB for the 4-kHz band.

A somewhat more demanding test of spatial uniformity was performed by exciting each loudspeaker sequentially with band-limited periodic noise, the spectrum of which consisted of lines of essentially equal amplitude spaced 10 Hz apart in frequency throughout the range 2 kHz to 6 kHz (the periodic noise output of the HP 3561A analyzer served as the basic source). This signal would excite specific standing waves in the room much more effectively than would the pink noise, which is random in character. Using the same sound level meter described above with the C-frequency-weighting network and "fast" exponential

averaging time, the difference between the highest and the lowest sound pressure levels measured throughout the listening area was 1.25 dB for each loudspeaker.

Care was taken to reduce the background noise levels in the listening room as much as practical. For example, the intake/exhaust air flows of the HVAC system for the room were balanced and unnecessary gratings in the air flow path were removed to reduce flow noise. The hot water flow to heat exchangers in the room was turned off during the listening tests and electrical resistance heaters were installed. The background noise level was then measured using a microphone system (B&K Type 4166 microphone, Type 2639 preamplifier, Type 2607 measuring amplifier), with the microphone at the listening position, and a dynamic signal analyzer (HP 3561A). The overall A-frequency-weighted sound pressure level was 32 to 33 dB. A one-third-octave-band analysis showed that most of this noise was contributed by frequency components below 250 Hz. Measurement of the unweighted one-third octave band levels showed that, for bands centered at frequencies greater than or equal to 750 Hz, the measured band levels were lower than the corresponding permissible one-third octave band levels [6] for ambient noise during audiometric testing using earphones mounted in the MX-41/AR cushion. For example, the permissible sound pressure level in the band centered at 4 kHz is 37 dB, and the corresponding level measured in the listening room is 6 dB. In other words, for frequencies near the encoding notch frequency, the listening room was quieter than an audiometric test booth suitable for determining audiometric thresholds of hearing with such earphones and cushions.

Measurements of the maximum C-weighted sound pressure level produced at the listening position by each of the musical selections presented in the serial listening tests were performed with the sound level meter set to the fast exponential averaging time. The highest such level was 88 dB; the lowest was 64 dB.

References for Appendix G

[1] Burkhard M. D., Sachs R. M. Anthropometric manikin for acoustic research. J. Acoust. Soc. Am. 1975; 58:219-22

[2] Burkhard M. D., ed. Manikin measurements. Elk Grove Village: Industrial Research Products, Inc., 1978.

[3] American National Standard For an Occluded Ear Simulator, ANSI S3.25-1979 (ASA 39-1979), American Institute of Physics for the Acoustical Society of America, New York.

[4] Atkinson E. S., Wahl G. L. Jr., Hall M. L., Wicklund E. J., Peterson S. K. Low-frequency analyzer combines measurement capability with modeling and analysis tools. Hewlett-Packard Journal 1987; 38 No.1:4-16.

[5] Epstein J. S. Versatile instrument simplifies dynamic signal analysis at low frequencies. Hewlett-Packard Journal 1984; 35 No. 12:4-11.

[6] Criteria for Permissible Ambient Noise during Audiometric Testing, ANSI S3.1-1977, American Institute of Physics for the Acoustical Society of America, New York.

APPENDIX H

Listener Instructions

APPENDIX H: LISTENER INSTRUCTIONS

Serial Presentation (presented aurally)

Thank you for your cooperation and your willingness to participate in these tests. We hope that you will find the tests both enjoyable and instructive. We realize that you are giving up prime family and work time to take part in the tests. We truly appreciate your contribution.

In order to ensure impartiality of the results, your experimenter is ignorant of the test key. When the tests are completed, we will break the code and notify you about your performance and that of the entire group. To obtain this report, please fill out your mailing address on the blank card provided at the end of the testing session.

Now please enter your name, the number of your listening station, and the date and time on your answer sheet. [wait 20 seconds]

For these tests, we either directly copied recorded music -- or we copied recordings through an Encoder which may alter signals in the frequency range near 3840 hertz which, musically speaking, falls between B-flat$_7$ and B$_7$. The frequency near 3840 hertz is near the top end of the piano and sounds like this. [wait 2 seconds] This frequency is above the fundamental frequency of most, but not all, musical instruments. Therefore, the Encoder will primarily operate upon the harmonics of lower fundamental frequencies, and thus may modify the timbre or coloration of the music.

In any given announced test, the same musical selection will be presented in pairs, with the two members of each pair either being the <u>Same</u> (that is, both direct or both encoded) or being <u>Different</u> (that is, one member of the pair being a direct reproduction and the other member being encoded). Each of the pairings will be repeated three times. You will then be asked to respond by encircling either <u>S</u>, for <u>Same</u>, or <u>D</u>, for <u>Different</u>, on your answer sheet. If you wish to change your answer, either erase or cross out your initial response and respond anew.

We will now run some practice pairs to acquaint you with the Encoder and the test procedure. The differences in the practice pairs may be more obvious than the differences within the later test pairs. Hopefully, the practice pairs will assist you in orienting your listening.

Here is practice pair <u>A</u> [wait 2 seconds]. Encircle either S or D in slot <u>A</u> depending upon whether you think the two members of the pair are the same or different. [wait 5 seconds].

[Repeat the above for practice pairs B, C, and D.]

The test administrator will now stop the tape to afford you an opportunity to ask any questions you may have about the test procedure.

We now will begin the actual testing. Please be ready to listen to the first test pair.

Here is test pair _1_ [wait 2 seconds]. Encircle in slot 1 either _S_ or _D_, depending upon whether you think the two members of the pair are the same or different. [wait 5 seconds]

Here is test pair _2_ [wait 2 seconds]. Encircle in slot 2 either _S_ or _D_.

[Continued for all 96 test pairs, with three short breaks.]

Parallel Presentation (read by subjects)

Thank you for your willingness to serve as a listener in this study. You have been asked to participate as a result of your excellent background in music and/or audio recording. We welcome and respect your expertise in careful and critical listening.

You have probably read in the newspapers or in the technical press that the Congress of the United States is considering legislation concerning a system that is intended to prevent the unauthorized copying of audio materials and that Congress has asked the National Bureau of Standards (NBS) to evaluate this system. This system, otherwise referred to as the "Copy-Code" system, consists essentially of two parts: (1) an Encoder designed to encode the two analog signals of a stereo recording by eliminating a small portion of the audio spectrum, thereby creating a "notch" in the recorded material, and (2) a detector that is designed to detect the presence of the "notch" in encoded material. The purpose of this listening study in which you are participating is to try to determine if the encoding process makes an audible difference in recorded material and, if so, to what degree.

You will be asked to listen to ten musical selections that were transferred digitally from commercially available compact discs to two tracks of a multi-track digital tape recorder. These two tracks were then copied on to six other tracks -- two of which constitute a reference (REF) of non-encoded (direct) material and four of which comprise two groups (A and B) of two tracks each, where either of the groups, A or B, may be direct or encoded material. Your task is to determine for each of the selections which group, A or B, is encoded material, i.e., is different from the reference track and the other group. The material in the given groups remains either direct or encoded throughout an entire selection but is subject to change from selection to selection in a random fashion. Further, for any given selection the material contained in the two groups A and B is always different, i.e., if for any given selection one group is encoded material, the other group is always direct material. You have been provided with a set of switches that enable you to switch between any of the groups of tracks: REF, A, and B. Each of the selections is preceded and followed by an announcement identifying the beginning and end, respectively, of the selection. The announcements are never encoded in either group A or group B regardless of the nature of the material preceding or following the announcements. The following is a table of example selections giving the nature of material in the three groups of tracks and the correct response relating to them.

Selection No.	A	Switch Position REF	B	Correct Response
N	Encoded	Direct	Direct	A
N+1	Direct	Direct	Encoded	B

Please circle the appropriate response on your answer sheet for each of the ten musical selections.

In order to preserve the levels of the original recordings, no attempt was made to compress the overall dynamic range contained in the recordings of the selected material. As a result, the maximum levels from selection to selection vary over a range of approximately 30 to 35 decibels. In order to provide you with adequate volume to listen to the selections recorded at relatively low levels, it was necessary that the selections recorded at relatively high levels be loud given the same relative volume control setting. There are two volume controls, one on the headphone amplifier, and one in a separate box adjacent to it. Please do not touch the volume control on the headphone amplifier, but do adjust the level of the recorded material to your liking using the attenuator in the separate chassis box. Be careful, however, to return this control to mid-position between selections, particularly after listening to a low-level recording. If possible, we ask that you record the setting of the attenuator during the time period that you evaluated each of the selections so that we have an idea of the nominal sound pressure level at which you listened to the recordings. The following is a list of the ten selections and the times, in minutes and seconds, at which they begin on the tape. Also noted is whether the level at which the material was recorded is relatively high or low (loud, soft, or average).

Selection No.	Level	Description	Tape Time
1	Average	Bizet-Sarasete: Carmen Fantasy	4:59
2	Loud	Streisand: Broadway Album	6:05
3	Soft	Beethoven: Symphony No. 9, Op. 125	7:55
4	Loud	Copeland: Fanfare for the Common Man	8:51
5	Loud	Pat Metheny: American Garage	10:11
6	Loud	Prokofiev: Alexander Nevsky Op. 78	10:52
7	Soft	Berstein: West Side Story	11:40
8	Average	Messian: Turangalila Symphony	12:55
9	Loud	J. S. Bach: Inventionen und Sinfonien	13:49
10	Soft	Respighi: Fountains of Rome	14:39

To enable you to listen to the recordings, you have been provided with a set of headphones. The acoustic frequency response of these headphones is dependent on the position of the phones on your head, particularly in the frequency region of 4000 Hz. Based on measurements made at NBS, we recommend that the headphones be positioned with the inner edge of the front of the ear cushions next to the pinnas of your ears, i.e., that the headphones be positioned towards the rear of your head. Also, be sure that the phones designated by "L" and "R" cover your left and right ears, respectively.

To assist you in evaluating the recorded material, you have been provided with a remote control unit to operate the tape recorder. This unit is capable of a number of functions, the operational details of which will be provided to you by the test administrator. A brief description of some of the features of the remote control unit follows. The remote control has two time displays: TAPE TIME, which indicates the actual tape position, and LOCATE TIME, which indicates the current time data to be used for a number of different functions as a destination address. These times may be used in either an absolute or relative mode with the mode indicated by the ABS TIME key/indicator. The tape times given for the selections are absolute and it will probably be less confusing to use the remote control in this time mode. The unit is capable of storing and recalling a total of 100 cue points which can be recalled and used as destination addresses for automatic operation of the tape transport. Cue registers 01 thru 10 contain the start times of selections 1 thru 10, respectively, as given in the table above. Destination addresses, or cue points, may be entered into a cue register either by using the numeric key pad and the STO key or by using the CUE STO key which stores the TAPE TIME at the time at which the key was pressed even if the transport is in the PLAY mode. Caution: when the CUE STO key is pressed the current cue register is incremented by one and the cue point in this register is overwritten by the new cue point. The cue register numbers are displayed in the CUE displays located next to the TAPE and LOCATE TIME displays. LOCATE TIMES, or destination addresses, may be expressed to the nearest second, or nearest millisecond, by pressing the mS/S key. Data may be entered into the LOCATE TIME display by recalling cue register contents, by pressing the CUE STO key, or by pressing the down shift key which transfers the data in the TAPE TIME display to the LOCATE TIME display. Three functions may be of use to you in operating the tape transport. Pressing the LOC key starts the locate operation which moves the tape to the destination address indicated by the LOCATE TIME display. Pressing the RLB key starts the roll back operation which moves the tape to the destination address indicated by the LOCATE TIME display less a preset PREROLL time (the default PREROLL time is two seconds but is easily changed using the numeric key pad and the STO and PRE keys). If the PLAY key is pressed at the same time as the LOC or RLB key is pressed, playback will automatically begin when the tape is located at the point defined by the operation. The RTN key starts the return, or loop, operation which moves the tape to a preset start point and then sets the transport into the PLAY mode until reaching a preset end point when it returns the tape to the start point and stops the transport (the start and end points are preset using the REP SET key to indirectly define destination addresses by cue register numbers which appear in the two CUE displays next to the time displays). If the PLAY key is pressed at the same time as the RTN key is pressed, the transport enters the REPEAT mode during which the transport continuously repeats playback of the tape between the start and end points. (Press one of the transport keys to stop a REPEAT operation.) Note: for your information the LEDs below the display should indicate the following conditions during playback: Fs: 44.1 kHz; STATUS: VARI-SPEED - off, MASTER SAFE -on, ALARM - off (except during selection no. 1 due to the emphasis status of the final recording differing from that of the original compact disc); REC MODE: INSERT. In general, you should ignore the status of these indicators and concentrate on listening to the recorded material.

We also are interested in receiving your comments on how you thought the encoded and direct material differed for each of the selections. A dictating machine is available for you to record any such comments that you wish to make as well as to record specific locations on the tape (TAPE or LOCATE TIMES) where the action of the encoding process was most evident to you.

You may wish to use the following information as a guide to aid you in ascertaining differences in the direct versus encoded recordings. The Encoder used in encoding the selections contains two narrow band-reject, or band-stop, filters each of which is simultaneously introduced into the analog signal path of one of the two channels of a stereo recording (with the filters assigned to separate channels). The filters are centered nominally about the frequency of 3840 Hz and are dynamic in the sense that the Encoder may activate or deactivate the filters in such a way as to either introduce them into, or remove them from, the signal paths of the two audio channels being encoded. The frequency of 3840 Hz is above the fundamental frequency range of most, but not all (e.g., the piccolo), musical instruments. Consequently, during most of the passages of encoded material, the filters of the Encoder, when introduced into the signal path, will affect the harmonics or overtones of lower fundamentals such as those produced by the human voice and by instruments such as the violin, trumpet, triangle, and bell. In the case of percussive instruments that have a broad spectral content such as the snare drum and cymbal, the apparent timbre, pitch, or loudness of an instrument may be altered by the encoding process. Also, the material that has passed through the Encoder, i.e., the "encoded material," may have been perceptibly altered by the Encoder during the time periods in the encoding process when the Encoder was introducing the filters into, or removing the filters from, the signal paths of the two audio channels. In any case, the differences in the encoded and direct material tend to be subtle rather than blatant and are likely to require careful, concentrated listening to detect differences between the two (for some of the selections it may be very difficult, if not impossible, to detect differences). A spectral analysis of the recorded material was performed after the recordings were completed and in all cases the analysis of the encoded material indicated that the filters of the Encoder were present in the signal path of the two audio channels for at least a portion of the time during the encoding of each of the selections. Since the recordings were made in such a way as to ensure that no perceptible differences were introduced into the direct versus encoded material by the recording process itself, you should concentrate on listening for differences between the direct and encoded material due to the filtering action of the Encoder rather than for differences in the material due to such things as variations in fade in and fade out of selections or low-frequency noise, relative time delays or lack of synchronization, etc. Also, we recommend for each of the selections that you listen to the entire selection that you are evaluating on each of the groups of tracks, REF, A, and B, before attempting to discern differences in the material by switching between them. (Note: a slight switching transient or "click" may be heard when switching between positions A, Ref, and B. This is an artifact of switching the digital audio output of the tape recorder and not due to differences in the recorded material).

APPENDIX I

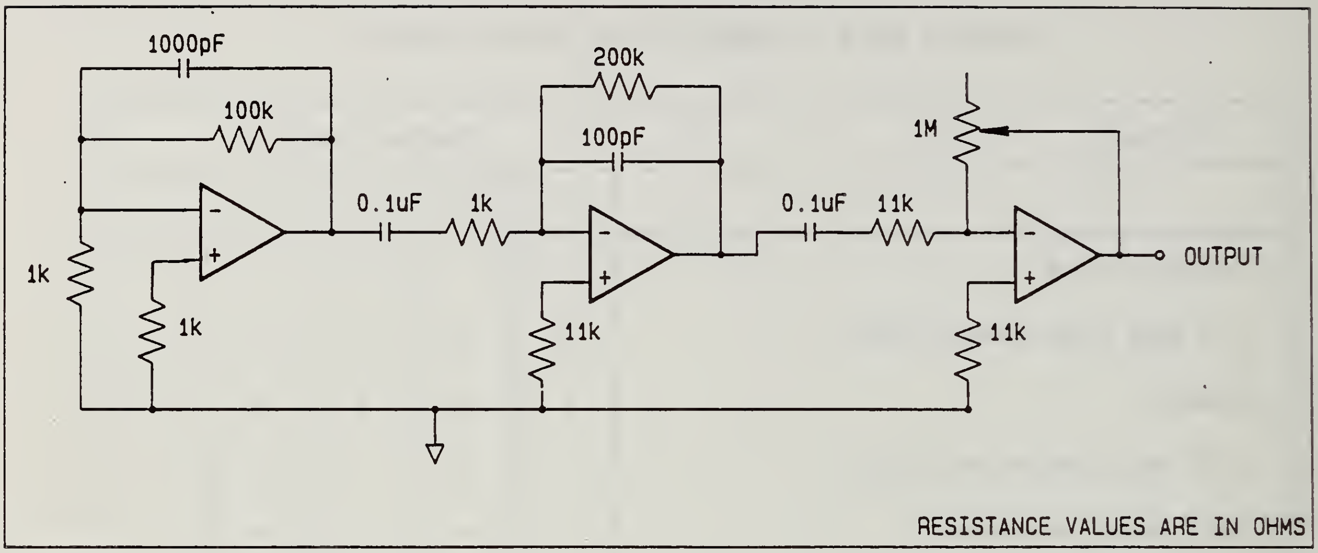
Schematic Diagrams of Defeat Circuits

Table I-1

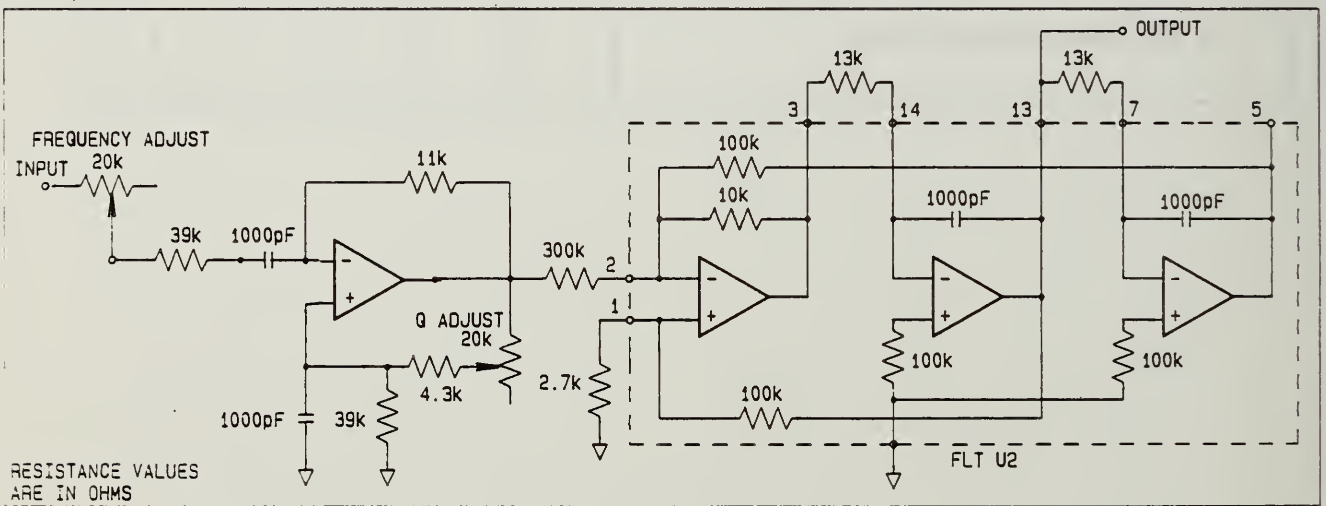
Circuits Used in Each of the Defeat Methods

Circuit	Defeat Method				
	1	2	3	4	5
Noise Source	•		•		
3.8 kHz Band Pass Filter	•	•	•	•	•
Summer	•	•	•	•	•(2)
1.9 kHz Band Pass Filter		•		•	•
Full Wave Rectifier		•		•	•
2.7 kHz Band Pass Filter			•	•	
Modulator/Multiplier/VC Amplifier			•	•	•
Full Wave Rect. and LP Filter			•	•	•
3.5 kHz Band Pass Filter					•
4.3 kHz Band Pass Filter					•

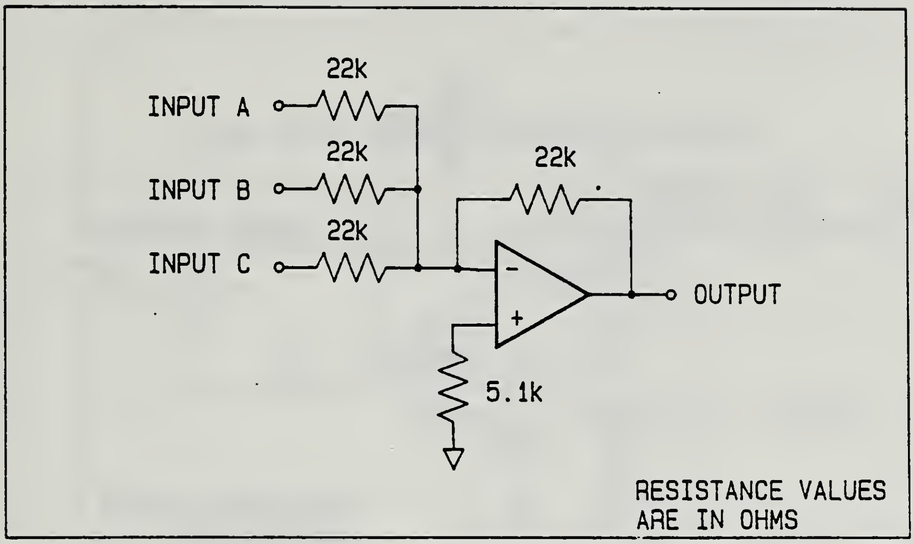
NOISE SOURCE



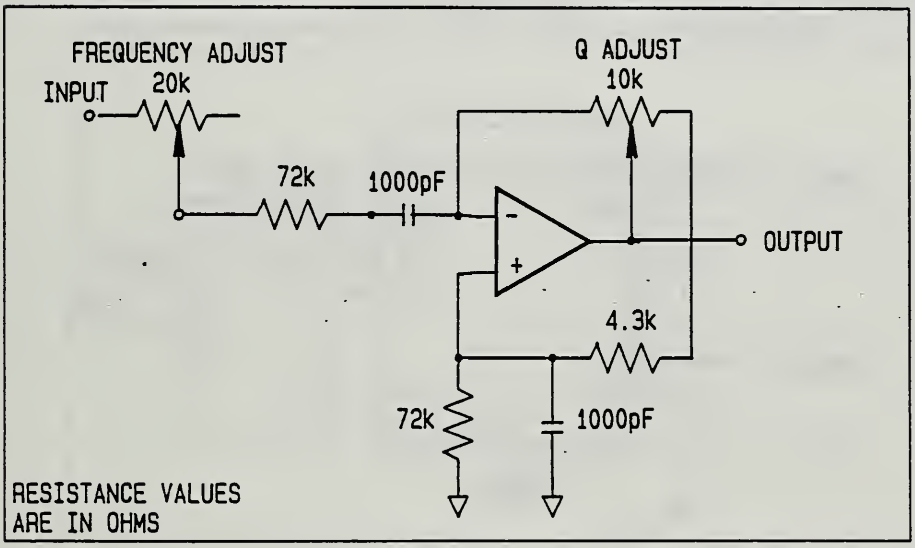
3.8 kHz BAND PASS FILTER



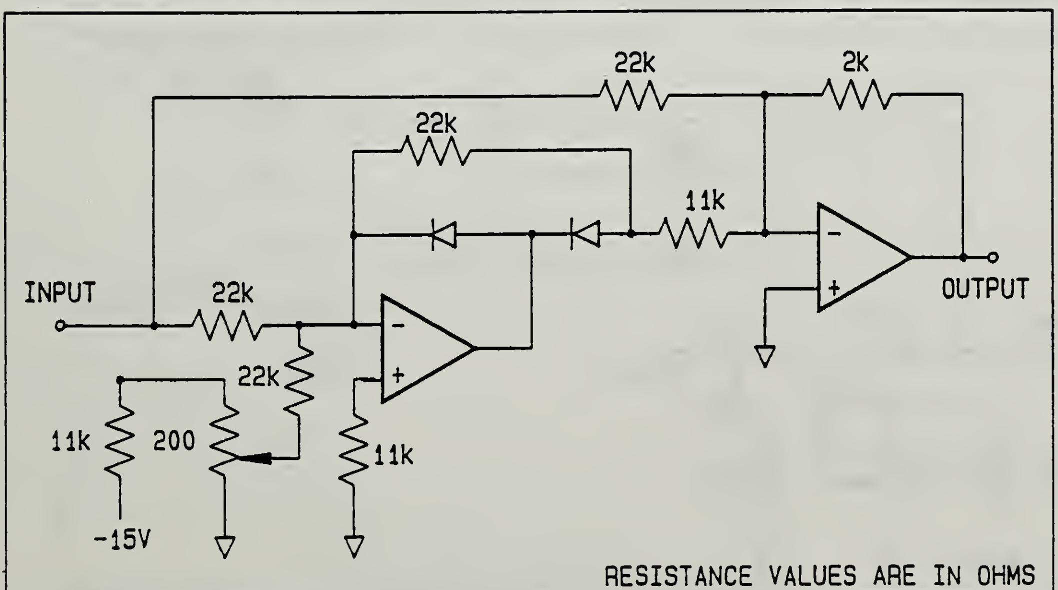
SUMMER



1.9 kHz BAND PASS FILTER



FULL WAVE RECTIFIER



NBS-114A (REV. 2-8C)

U.S. DEPT. OF COMM. BIBLIOGRAPHIC DATA SHEET (See instructions)	1. PUBLICATION OR REPORT NO. NBSIR-88/3725	2. Performing Organ. Report No.	3. Publication Date February 1988

4. TITLE AND SUBTITLE

Evaluation of a Copy Prevention Method for Digital Audio Tape Systems

5. AUTHOR(S) B.A. Bell and G.N. Stenbakken, D.R. Flynn, D.J. Evans, E.D. Burnett, V. Nedzelnitsky, and K.R. Eberhardt

6. PERFORMING ORGANIZATION (If joint or other than NBS, see instructions)

NATIONAL BUREAU OF STANDARDS
DEPARTMENT OF COMMERCE
WASHINGTON, D.C. 20234

7. Contract/Grant No.

8. Type of Report & Period Covered

9. SPONSORING ORGANIZATION NAME AND COMPLETE ADDRESS (Street, City, State, ZIP)

10. SUPPLEMENTARY NOTES

☐ Document describes a computer program; SF-185, FIPS Software Summary, is attached.

11. ABSTRACT (A 200-word or less factual summary of most significant information. If document includes a significant bibliography or literature survey, mention it here)

The National Bureau of Standards in response to requests from the U.S. Congress tested a system designed to prevent unauthorized copying by digital audio tape (DAT) recorders of suitably encoded audio recordings. The system, designed by CBS Records, filters out a narrow range of frequencies from the spectrum of the original sound in the region of 3840 Hz, thereby encoding the material with a "notch "in the frequency spectrum so that a DAT recorder equipped with the system's decoding circuitry can sense the presence of a prescribed notch in the spectrum and inhibit recording.

The congressional questions and the NBS conclusions are:

1. Does the copy prevention system achieve its purpose?

NBS Conclusion: The system does not achieve its stated purpose.

2. Does the system diminish the quality of the prerecorded material into which the notch is inserted?

NBS Conclusion: The system's encoder alters the original electrical signal. For some listeners for some selections, this results in a discernible difference between prerecorded notched and unnotched material.

3. Can the system be bypassed and, if so, how easily?

NBS Conclusion: The copy prevention system can be bypassed easily.

12. KEY WORDS (Six to twelve entries; alphabetical order; capitalize only proper names; and separate key words by semicolons)

audio recording; copy protection; digital recording; music; sound; tape recording

13. AVAILABILITY

☒ Unlimited

☐ For Official Distribution. Do Not Release to NTIS

☐ Order From Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402.

☒ Order From National Technical Information Service (NTIS), Springfield, VA. 22161

14. NO. OF PRINTED PAGES 205

15. Price $24.95